

A Unified Theory of Dimensional Aesthetics

Toward Nervous-System-Compatible Digital Environments

Revised Draft v5 — Research Program Edition, with the Formative Corpus

Author: Olya Androsik

Affiliation: Independent Researcher; Design Systems & Cognitive Safety, Auro Prism

21 October 2026

Draft for review. This revision (v5, 21 October 2026) follows the Research Program Edition (v4, 1 October 2026), which answered a structured multi-perspective review of Draft v3 (August 2026). v5 changes no construct, no claim tier, and no checkable form. It does two things: it declares the author's formative corpus as research material, and it rewrites the prose for precision. A summary of revisions precedes the paper; the Canon Edition addendum follows it.

Summary of Revisions from v4

v4 rebuilt the framework's architecture in answer to review. v5 leaves that architecture untouched and does two things the review did not ask for and the author owes.

1. The formative corpus is declared. The constructs in this paper were not derived from the cited literatures alone. They were formed by a reader — in Russian before English — of a particular set of novels, plays, poems, and songs, and by training in interactive, sound, and game arts. v4 named "design practice" and "lived experience" as sources and left the rest undeclared. A framework whose audit requires every friction point to carry a legible purpose should not keep its own sources illegible. §2.11 declares the corpus and fixes its epistemic role: a source of constructs, questions, and the standard of description, never a source of evidence. Appendix D registers each work against the construct it shaped. Lineage markers appear in the body where a construct has a literary ancestor (§2.2, §2.4, §3.0, §3.1, §3.3, §3.5, §3.6, §4 Doctrines IV, IX, and XI, §5.3, §5.7, §6.7, §7, §8, §9). Nothing at Tier A or Tier C rests on any of them.
2. The prose is rewritten for precision. The v4 text was accurate and heavy: sentences stacked on semicolons, categories where instances were available, hedges repeated after the tier labels had already done their work. v5 shortens the sentences, names the thing rather than its class (a countdown on a checkout page, not "urgency mechanics"), and removes the hedges that the claim architecture (§1.3) already supplies. The rule of the Canon addendum governed the pass: prose may change in voice, never in claim strength. Where a v4 sentence carried a commitment, the commitment is in v5, shorter.
3. The positionality statement is extended with the standpoint facts the corpus depends on: two languages, two literatures, a move between them at sixteen, and an art-school training in interactive, sound, and game arts — the origin of this paper's habit of treating an interface as a timed, orchestrated environment rather than a surface.
4. Appendix D (the corpus register) and a References subsection, Literary and musical sources, are added. The Canon addendum gains one clause: the corpus is Canon material, and the Canon edition may use it more freely than the paper does.

Everything else — five dimensions, two cross-cutting properties, eleven doctrines with their checkable forms, the correspondence table, four claim tiers, the six-study program, the Sensing Wall, the limitations — is carried from v4 with changes of wording only.

0. Terminology Note

Nervous-system-compatible names this framework's ambition: digital environments whose demands stay within, and respect, human regulatory capacity. The phrase is kept as a vision term. Its technical content here is narrower than the phrase suggests, and four disclaimers bound it.

First, no claim in this paper depends on a universal stress response, a universal calming stimulus, or a fixed mapping from interface properties to emotional states. The affective science indicates that no such mappings exist (Barrett, 2017; Barrett et al., 2019). Second, no claim relies on polyvagal theory, which is contested and is not cited as a mechanism. Third, where the paper speaks of physiological cost, the construct is allostatic load — the cumulative cost of regulation under demand, in the predictive-regulation sense of Sterling (2012) and McEwen — advanced as a labeled hypothesis (§7), not as established interface science. Fourth, the operational claims — the dimensions, the doctrine, the audit — are stated and testable in behavioral and self-report terms. Withdraw the physiological hypothesis entirely and the design and governance claims stand unchanged. A reader who translates "regulatory burden" as "demand on attention, comprehension, decision, and recovery, measured behaviorally" loses nothing operational.

Author Positionality and Intent

This framework comes from professional practice in product design, design systems, and design operations; from neurodivergent lived experience; and from work alongside therapy-adjacent contexts. It does not prescribe clinical practice and makes no diagnostic claim. Its intent is to define environmental and interactional constraints that may reduce avoidable cognitive, sensory, and emotional harm in digital systems.

One clarification is structural. Lived experience informs this framework; it cannot legitimate it. One insider is a participation of exactly one, and a baseline defined by one author would contradict its own pluralism. The claims this paper makes on behalf of neurodivergent users are propositions for community validation, not settled requirements, and the research program (§6.7) binds their development to participatory methods with constitutive community power. Where this paper and future community review conflict, the community review governs.

A second clarification is new to this revision. The author grew up in Minsk reading in Russian, moved to the United States at sixteen and read the English canon there, and trained at art school in interactive arts, sound art, and game arts before working in product design. Those facts are not biography for its own sake; they are the provenance of the paper's instruments. The habit of treating an interface as a timed, orchestrated environment comes from sound and

games, where a designer sets a rhythm and a demand curve deliberately. The habit of asking what a designed environment does to the people inside it, and of describing it exactly, comes from a particular set of books and songs. §2.11 declares that corpus and bounds what it is allowed to do in a scientific paper. The declaration follows the paper's own rule for interfaces: a source that does work should be visible, and its purpose legible.

Abstract

Digital environments set the conditions under which people attend, decide, regulate, trust, and act. Interfaces mediate healthcare, therapy, learning, finance, work, identity, social participation, and AI-assisted decision. In a development that postdates most of the relevant design literature, regulators now treat interface mechanics as objects of enforcement: cancellation-flow architecture, feed rhythm, autoplay, notification cadence, and the dismissibility of protective features appear in settlements, statutory rules, and platform-liability findings. Design practice still evaluates digital systems mainly by usability, efficiency, and engagement — metrics that capture a narrow part of what interfaces do to the people using them.

This paper presents the Unified Theory of Dimensional Aesthetics (UTDA): a theoretical synthesis and a design research program for evaluating digital interfaces as multi-dimensional cognitive and affective environments. "Aesthetics" here is not visual style; it is the whole perceptual, temporal, semantic, and interactional structure through which an interface is experienced. The framework identifies five stimulus-side dimensions — visual density, motion energy, temporal rhythm, interaction depth, and semantic load — each anchored to prior validated constructs and candidate measures, and two cross-cutting properties that emerge from their configuration: predictability, divided into veridical and experienced forms, and affective affordance, which replaces v3's "emotional valence." Its central empirical claim is falsifiable: dimensional coherence — the degree to which dimensions reinforce rather than contradict one another — predicts user strain beyond the intensity of any single dimension.

On that model, the Auro Prism Doctrine states operational principles for design evaluation and governance: coherence as a structural property, veridical predictability before personalization, regulation before engagement, legible friction, motion as a cognitive instrument, proportionate semantic load, neurodivergent safety as a co-produced baseline, transparent adaptation, dimensional and ethical bounds on AI-generated interfaces, structural ethics, and one prohibition — the Sensing Wall — against turning dimensional evaluation into user-state sensing.

The paper's claims are stratified into four labeled tiers: behavioral-normative claims defensible on current evidence; theoretical synthesis; a quarantined physiological hypothesis (allostatic cost) with stated falsification conditions; and operational proposals whose adoption is conditioned on a six-study validation program in dependency order, beginning with inter-rater reliability. The framework prescribes parameterized pluralism — population-calibrated dimensional profiles — rather than universal thresholds, binds its neurodivergence claims to participatory governance, and, in this revision, declares the literary and musical corpus from which its constructs were formed, with that corpus bounded to the role of source, never evidence.

Keywords: dimensional aesthetics, dimensional coherence, regulatory burden, allostatic load, cognitive load, neurodivergent-inclusive design, interface ethics, dark patterns, deceptive design, humane technology, generative interfaces, AI governance, appropriate reliance, research through design

1. Introduction

1.1 The problem, and why now

Human-computer interaction has long addressed usability, accessibility, efficiency, learnability, and satisfaction. Later traditions — calm technology, trauma-informed computing, dark patterns research, design justice, digital wellbeing, human-centered AI — widened the field's concern to harm, agency, inclusion, and manipulation (Weiser & Brown, 1996; Chen et al., 2022; Gray et al., 2018; Costanza-Chock, 2020; Amershi et al., 2019). Together they establish that an interface is not a neutral container. It decides what a user notices, how options read, how much effort an action costs, and how much strain the person absorbs on the system's behalf.

Two developments make the present different from the moment in which those literatures formed.

The first is regulatory. Between early 2025 and mid-2026 the object of technology regulation moved from data practices to design mechanics. The European Union's prohibition on manipulative and exploitative AI techniques became applicable law (EU AI Act, Art. 5, in force February 2025). The U.S. Federal Trade Commission settled with Amazon for \$2.5 billion over a subscription-cancellation flow alleged to obstruct users through its architecture, with injunctive relief that mandates specific interface changes and independent compliance monitoring (FTC, 2025). The European Commission issued a preliminary finding that TikTok's design — infinite scroll, autoplay, notification cadence, screen-time tools that dismiss with one tap — itself breaches the Digital Services Act's obligations toward minors and vulnerable

users (European Commission, 2026). New York finalized rules restricting algorithmic “addictive feeds” for minors (SAFE for Kids Act rules, 2026), and the EU’s Digital Fairness Act, aimed at dark patterns and addictive design generally, entered the legislative calendar. Whatever one thinks of any single instrument, the pattern is plain: feed rhythm, cancellation depth, motion, interruption timing, and friction asymmetry are now legally cognizable properties of a design. Regulators are enforcing dimensional constraints without a dimensional theory. Courts, meanwhile, have struck vague design mandates, which argues for precise, measurable frameworks over general wellbeing language.

The second is generative. Interfaces are increasingly assembled by AI systems at generation time — per user, per session, at a speed and scale beyond human review. In two years the industry conversation moved from demonstrations to documented quality anxieties: drift, incoherence, inconsistent tone, interface states nobody reviewed. Existing human-AI interaction guidelines govern interaction behavior — expectations, errors, control (Amershi et al., 2019; Shneiderman, 2022). No published framework governs the sensory, temporal, and semantic structure of a generated interface. The window in which such a framework can shape practice is the window in which generative interfaces are being made into infrastructure, which is now.

Against this background the gap is the one earlier drafts named, now with outside corroboration: designers, organizations, and regulators lack a shared operational language for how the sensory, temporal, semantic, and interactional properties of an interface combine to act on the person using it. Accessibility standards identify barriers; usability methods detect task friction; dark patterns research names manipulation; trauma-informed computing surfaces safety; calm technology argues for respectful attention. None of them — and enforcement actions are already reaching for it — provides a dimensional model of the interface as one combined environment.

1.2 About Auro Prism

Auro Prism is a design initiative developing standards, tools, and applied methods for nervous-system-compatible digital environments (see §0 for the technical scope of the phrase). Its work spans theoretical research, design doctrine, interface evaluation, AI-assisted product governance, and applied design practice. This paper is the theoretical foundation and research program for that work.

1.3 Claim architecture

Every claim in this paper belongs to one of four tiers, and the tier is load-bearing: readers, reviewers, and adopters should hold each claim to the standard of its tier and to no other. The stratification answers a hazard the paper takes seriously — in synthetic frameworks,

speculative claims borrow credibility from defensible neighbors.

Tier A — Behavioral-normative claims. Claims about what interfaces observably do and what they ought not do, defensible on current behavioral evidence, existing literatures, and enforcement records. Example: a cancellation flow that adds asymmetric effort, hides the exit, and uses guilt-laden copy obstructs the user's intent; this is measurable behaviorally and has already been priced at \$2.5 billion. Nothing in Tier A requires physiology.

Tier B — Theoretical synthesis. The dimensional taxonomy and the coherence construct as an analytic structure integrating prior literatures. Tier B claims are judged by fruitfulness and by correspondence to prior constructs (§2.9), pending the psychometric program (§6) that tests whether the structure survives contact with data. The taxonomy is provisional: if factor-analytic derivation restructures it, the framework versions accordingly.

Tier C — The physiological hypothesis. The proposition that dimensional demand levies allostatic cost — that regulatory burden has a bodily accounting beyond momentary task load. Tier C is quarantined in §7, carries falsification conditions, and is not a premise of the doctrine or the audit. It is the framework's research frontier, not its foundation.

Tier D — Operational proposals. The doctrine, state classification, review heuristics, and audit machinery. Their authority is conditional: on inter-rater reliability being demonstrated (§6.1), on participatory validation for claims made on behalf of neurodivergent users (§6.7), and on the anti-gaming provisions of §6.8. Until those conditions are met, Tier D is a structured proposal for review practice, not a standard.

The formative corpus (§2.11) sits outside the tiers. It supplies constructs, questions, and language; it is cited only where a construct's ancestry is being stated, and a literary citation in this paper is never a citation of evidence.

1.4 Contributions

This paper makes seven contributions. (1) It introduces nervous-system compatibility — technically, regulatory burden and its allostatic hypothesis — as a design and governance concern, stratified into explicit claim tiers. (2) It proposes the Unified Theory of Dimensional Aesthetics: five stimulus-side dimensions anchored to prior constructs and candidate measures, with a correspondence table to existing literatures. (3) It states the framework's distinctive falsifiable prediction — coherence predicts strain beyond single-dimension intensity — and designs the study that would test it. (4) It reframes predictability (veridical versus experienced) and affect (affective affordance) in forms consistent with current affective science and appropriate-reliance research, closing the loophole ("manufactured calm") that an earlier draft left open. (5) It revises the Auro Prism Doctrine into behaviorally checkable form, adds the Sensing Wall, binds neurodivergent-safety claims to participatory

governance, and maps the doctrine's failure modes onto the dark-patterns ontology rather than beside it. (6) It presents a sequenced, dependency-ordered validation program with anti-gaming design, cross-cultural calibration, and a defined division between behavioral evidence (near term) and physiological evidence (long term, behind the Sensing Wall). (7) It declares the author's formative corpus as research material with a bounded epistemic role, so that the framework's provenance is as legible as the interfaces it proposes to audit.

2. Related Work and Theoretical Foundations

The framework does not replace prior traditions; it integrates them through a dimensional model centered on cognitive-affective demand and operational design constraints. This section reviews the foundations, adds the literatures earlier drafts lacked, closes with the two structural devices introduced in v4 — a correspondence table (§2.9) anchoring each dimension to prior constructs, and a mapping onto the dark-patterns ontology (§2.10) — and ends with a section new to this revision, the declared corpus of literary and musical sources (§2.11).

2.1 Cognitive Load Theory

Cognitive Load Theory grounds the claim that informational environments tax working memory (Sweller, van Merriënboer, & Paas, 2019). In an interface the load comes not only from content but from navigation structure, visual density, interaction sequence, terminology, timing, and interruption. The usual UX response is reduction — fewer words, options, steps — and reduction can fail: simplicity that is not signaled raises interpretive effort, because the user must infer the relationships and consequences the page no longer shows. Hiding structure does not remove information; it moves the processing from the page to the person. Two cautions from the measurement literature are adopted. Subjective load instruments (Paas-type scales, NASA-TLX) face unresolved validity debates, and the separability of load types is disputed; any dimension that inherits cognitive-load constructs inherits this unfinished measurement problem, which is why the validation program starts with reliability rather than theory (§6.1). And capacity is not a fixed count: it depends on chunking, expertise, and meaning (Miller, 1956) — a standing argument against fixed numeric thresholds anywhere in this framework.

2.2 Prediction, Expectation, and Regulatory Cost

Earlier drafts used predictive processing (Friston, 2010; Clark, 2016) as a lens. This framework keeps the narrower design claim and drops any dependence on the contested theory. The design claim: an environment that forces continuous expectation revision without adequate

feedback, structure, or control imposes measurable cost. The cost is demonstrable through expectation-violation paradigms, uncertainty effects on decision performance, and vigilance costs. None of this commits the framework to free-energy formulations, whose status as neuroscience is still argued (Bruineberg et al., 2022; Gershman, 2019). An interface teaches a user what to expect through layout, hierarchy, timing, motion, feedback, and language. When those signals agree, the user anticipates and conserves effort; when they conflict, the user pays in vigilance and re-evaluation. An old experimental literature makes the same point from physiology: organisms required to sustain a discrimination under progressively contradictory signals break down systematically — Pavlov's "experimental neurosis" preparations (Pavlov, 1927) showed, nearly a century ago, that unresolvable signal conflict is not merely unpleasant but disorganizing. The coherence construct (§3.8) is the design-level descendant of that observation, advanced at Tier B. Its literary ancestor is declared in §2.11: the author met contradictory signal as a subject before meeting it as a finding, in a Russian literature that specializes in describing what a person does inside an environment whose rules keep changing.

2.3 HCI, Usability, and Situated Interaction

The foundational HCI concepts — affordance, feedback, visibility, mapping, consistency, error prevention, user control (Norman, 2013; ISO 9241-210:2019) — stay central. The framework's argument is that conventional evaluation is incomplete, not wrong: task completion does not show regulatory safety; speed does not show clarity; satisfaction does not reveal coercion or the depletion that arrives after the task. A regulatory approach adds four questions. Did the interface preserve orientation? Did it make change visible? Did it communicate friction honestly? Did it require vigilance it did not need? The same lineage supplies the framework's methodological conscience: the Model Human Processor and GOMS showed that interface experience becomes calculable when researchers commit to parameters and prediction (Card, Moran, & Newell, 1983). The framework accepts that as its own eventual bar and approaches it first through its best-parameterized dimension, temporal rhythm, where chronometric estimates already exist (§6.5).

2.4 Neurodivergence, Sensory Processing, and Interface Safety

Neurodivergent users are routinely treated as edge cases, accommodated after the primary experience has been defined. This framework rejects that sequence and holds neurodivergent safety to be a baseline condition (Milton, 2012; Spiel, Frauenberger, Keyes, & Fitzpatrick, 2019; Mack et al., 2021). The theoretical warrant is the double empathy problem (Milton, 2012; Milton, Gurbuz, & López, 2022): interaction failure between differently constituted parties is a mutual mismatch, not a deficit in one of them — moved here from the interpersonal to the human-interface dyad. The frame has teeth, and this framework accepts them. If mismatch is

mutual and perspectival, there is no neutral standpoint from which one evaluator defines “coherence.” A framework consistent with its own theory must therefore be dialogical: plural safety profiles, negotiated standards, evaluator panels whose disagreements are reported as data rather than averaged away (§6.7). The oldest case study of the problem in the author’s corpus is Dostoevsky’s Prince Myshkin (1869): a man who reads the people around him exactly, takes their words as meant, and is read by them as defective. The novel locates the failure in the room. Dick’s Voigt-Kampff test (1968) is the same case from the other side: an empathy test administered by one party classifies the other as defective, and the novel doubts its own instrument (§2.11). The baseline claim is correspondingly precise. Digital systems should not treat high sensory tolerance, rapid context switching, ambiguous hierarchy, forced urgency, dense language, continuous adaptation, or constant availability as neutral defaults; they are not neutral; they favor some processing profiles and tax others. A social argument reinforces the ethical one: when safety is the default, invoking it discloses nothing about the user. An accommodation mode that visibly simplifies a screen is a disclosure event in company — legible to onlookers, and so potentially stigmatizing — whereas baseline safety is invisible because universal.

2.5 Wellbeing-Supportive Design and Self-Determination Theory

A literature absent from earlier drafts anchors the framework’s psychological register. Positive computing made wellbeing a design criterion organized around psychological needs — autonomy, competence, relatedness — with validated instruments and a spheres-of-influence model separating interface-level from life-level effects (Calvo & Peters, 2014; Peters, Calvo, & Ryan, 2018). The present framework is complementary: where need-based models say what people require, the dimensional model says how stimulus and interaction structure tax or support the capacities those needs presuppose. Two adoptions follow. Perceived competence and autonomy instruments from this tradition serve as near-term outcome measures for dimensional studies (§6), which spares the framework the invention of unvalidated scales. And the construct nearest the framework’s “orientation” — the user’s stable sense of where they are, what changed, and what they can do — is aligned with self-efficacy (Bandura, 1997), which has five decades of measurement behind it. An interface that preserves orientation and legible recovery manufactures efficacy beliefs. A dysregulating interface’s deepest cost may be learned inefficacy: the accumulated lesson that one cannot understand or safely operate one’s own environment.

2.6 Attention, Interruption, and Digital Self-Control

Empirical attention research supplies the framework’s best-evidenced dimension. Field studies document steeply shortened screen-attention durations and large interruption-recovery costs, together with a rhythmic structure to attention — focus and recovery

alternate, and some interruptions serve wellbeing while others fragment it (Mark, 2023; Mark, Gudith, & Klocke, 2008). HCI research has characterized attention-capture damaging patterns — infinite scroll, autoplay, guilty-pleasure recommendations — with empirical methods (Monge Roffarello, Lukoff, & De Russis, 2023), produced designer-facing attention heuristics grounded in self-determination theory (Monge Roffarello, De Russis, & Lukoff, 2025), and meta-analyzed digital self-control tools (Lyngs et al., 2019). Two obligations follow. The temporal-rhythm audit must not score all interruption as harm, because protective interruptions exist on the record. And claims about attention are phrased as measured strain in studied populations, not as universal cognitive law.

2.7 Adjacent Frameworks

Calm Technology holds that technology should respect attention and move between periphery and center without intrusion (Weiser & Brown, 1996; Case, 2015); it has since become a certification standard, mainly for hardware (Calm Tech Institute, 2024). This framework extends the concern across dimensions and into software and generated interfaces, and adds the constructs — predictability, affective affordance, coherence — that one calm/attention axis cannot carry. Trauma-Informed Computing adapts safety, trust, and empowerment principles to computing (SAMHSA, 2014; Chen et al., 2022; Scott et al., 2023). Its “universal precaution” logic — design for the most affected and everyone benefits — is the nearest relative of this framework’s baseline claim. The two differ in kind: principles-based and trauma-centered against dimensional and demand-centered. Design Justice (Costanza-Chock, 2020) supplies the structural conscience: the question is never only what the default does but who defined it and who governs its evolution — an obligation this framework discharges structurally (§6.7) rather than rhetorically. Slow Technology (Hallnäs & Redström, 2001; Odom et al., 2012) and design frictions and microboundaries research (Cox et al., 2016) establish that added effort can serve reflection and self-control; the protective-friction construct is their systematized descendant, and the prior naming is acknowledged. Dark patterns research (Gray et al., 2018; Mathur et al., 2019; Mathur, Mayer, & Kshirsagar, 2021; Gray et al., 2024) is treated in §2.10. Digital wellbeing scholarship, including its critique of static conceptions (Vanden Abeele, 2021), warns against treating wellbeing as a fixed target — congruent with parameterized pluralism. Human-Centered AI (Amershi et al., 2019; Shneiderman, 2022; Weidinger et al., 2022) governs AI interaction behavior; this framework extends governance to the sensory and structural properties of AI-generated interfaces, where no published framework yet operates. Value Sensitive Design (Friedman & Hendry, 2019) contributes method rather than content: its stakeholder investigations, indirect stakeholders included, enter the audit program (§6.7) in preference to parallel machinery.

2.8 Aesthetics Measurement, Neuroaesthetics, and Affective Design

Three further literatures, absent from earlier drafts, are foundational to what this framework attempts. First, interface aesthetics measurement. Perceived visual aesthetics decomposes into replicable factors (classical and expressive aesthetics: Lavie & Tractinsky, 2004; simplicity, diversity, colorfulness, craftsmanship: Moshagen & Thielsch, 2010); first impressions form within tens of milliseconds and are predictable from computable properties including visual complexity (Reinecke et al., 2013); automated interface-aesthetics metrics exist (Miniukovich & De Angeli, 2015). This literature gives both precedent and method for dimensional decomposition of interface experience, and its psychometric pipeline — large stimulus sets, ratings at scale, factor derivation — is the method by which this framework's authored taxonomy must eventually be tested and, if necessary, restructured (§6.3). Second, environmental neuroaesthetics. The aesthetic triad model and its empirical program derived dimensions of architectural experience — coherence, fascination, hominess — by psychometric decomposition and connected them to affective response (Coburn, Vartanian, & Chatterjee, 2017; Coburn et al., 2020), while environmental psychophysiology has measured stress responses to built environments in the field (Ellard, 2015). This is the framework's closest methodological ancestor: environment as input to a regulating organism, evaluated dimensionally. The architectural "coherence" dimension names perceptual organization within a scene; this framework's coherence names alignment across stimulus dimensions — related, distinct, and now differentiated. Porting that research program from buildings to screens, where people now spend more waking hours than in any building, is the framework's central methodological proposal. Third, Kansei engineering (Nagamachi, 1995): a mature industrial method mapping product design parameters to multidimensional affective response, acknowledged as a fifty-year precedent for dimensional affective design, differing in aim (preference optimization) from this framework's aim (demand governance). Adjacent to all three stands game design, with decades of applied knowledge in orchestrating density, pacing, challenge, and reward on purpose. Flow theory and difficulty-curve practice (Csikszentmihalyi, 1990) are the nearest existing experimental corpus on deliberate dimensional orchestration, and they instruct in both directions: how demand builds capability, and how the same machinery becomes extraction. The author's training in game arts (see Author Positionality) is where this paper learned to read an interface as a demand curve.

2.9 Correspondence of Dimensions to Prior Constructs

The table is a structural commitment, not a courtesy: each dimension is anchored to prior constructs and candidate measures, and a dimension that cannot keep such an anchor does not belong in the framework.

UTDA construct	Nearest prior constructs	Existing measures / instruments	Status
Visual density	Visual complexity; clutter; classical aesthetics (inverse)	Computable complexity metrics (Reinecke et al., 2013; Miniukovich & De Angeli, 2015); VisAWI simplicity facet	Tier B; measurable now
Motion energy	Animation-from-interactions (WCAG 2.3.3); attentional capture by motion	Motion-capture attention paradigms (Franconeri & Simons, 2003); reduced-motion compliance checks	Tier B; partial measures
Temporal rhythm	Interruption cost; attention rhythms; system response time; calm-tech periphery	Chronometric parameters (Card et al., 1983); interruption/recovery field methods (Mark et al., 2008)	Tier B; best-parameterized
Interaction depth	Task steps (GOMS); attention-capture patterns; microboundaries; sense of agency	GOMS-style task analysis; attention-capture pattern coding (Monge Roffarello et al., 2023)	Tier B; measurable now
Semantic load	Intrinsic/extraneous load; linguistic surprisal; readability; COGA clear-language objectives	Load scales (with known caveats); psycholinguistic surprisal models (Levy, 2008); COGA checks	Tier B; measurement contested
Veridical predictability	Consistency; mental-model match; appropriate reliance	Expectation-violation rates; reliance-calibration tasks (Liao et al. tradition)	Tier B; measurable now
Affective affordance	Appraisal affordances; hedonic quality; emotional design (visceral level)	Appraisal-based self-report; explicitly NOT emotion detection	Tier B; reframed this draft
Dimensional coherence	Gestalt unity-in-variety; architectural coherence (distinct); signal conflict (Pavlov, 1927)	None yet — the framework's distinctive construct; operationalized §3.8, tested §6.2	Tier B; the flagship claim
Regulatory burden	Workload; ego-depletion (rejected: see §3.11); allostatic load (Tier C)	Behavioral battery (Appendix C); ambulatory physiology (Tier C only)	A (behavioral) / C (physiological)

2.10 Relation to the Dark Patterns Ontology

After years of parallel taxonomies, deceptive-design research has consolidated into an ontology of high-, meso-, and low-level patterns (Gray et al., 2024) that increasingly anchors shared regulatory language. This framework does not introduce a rival taxonomy of manipulation. Its failure modes map onto the ontology: friction failure corresponds to the ontology's obstruction strategies; semantic failure spans sneaking and interface interference; regulatory failure covers forced action and social-engineering clusters seen through their demand on the user; predictability failure and orientation failure mostly concern interface interference and sneaking as experienced over time. What the framework adds above the ontology is an explanatory and evaluative layer: an account of why these patterns harm (they levy demand and distort expectation) and the constructive counterpart the deceptive-design literature has repeatedly asked for and not systematized — a specification of protective structure: legible friction, transparent adaptation, proportionate load, coherent signaling. In a regulatory setting the ontology names the violation; this framework explains its cost and specifies its remedy.

2.11 The Formative Corpus: Sources Outside the Scientific Literature

A framework is also a reading. The constructs in §3 were authored before they were anchored (§3.0 says so), and the author who wrote them had been formed by particular books and songs. This section declares them. The warrant is methodological. Research through design treats the designer's trained judgment as an instrument of inquiry (Zimmerman, Forlizzi, & Evenson, 2007; Bardzell, Bardzell, & Hansen, 2015), and an instrument is declared and calibrated, not hidden. §6.1 calibrates the reviewers' judgment against one another; this section calibrates the author's against its sources. It also follows the paper's own rule for interfaces. Doctrine IV requires that every friction point carry a legible purpose and Doctrine VIII that a system remain transparent about why it behaves as it does. A paper that imposes those rules on a checkout flow should meet them itself.

Epistemic role. The corpus is bounded to one job. It supplied constructs, questions, names, and a standard of description. It supplies no evidence. No Tier A claim cites a novel; no Tier C claim cites a poem; no checkable form in §4 contains literary content; no study in §6 depends on a reading. Where a literary work is cited in the body, the citation marks ancestry — this construct was first met here — and nothing else. A reader who deletes every literary reference from this paper loses provenance and loses no claim.

What the corpus is. Two literatures, read in sequence. The Russian one came first and in the original: Pushkin, Gogol, Dostoevsky, Tolstoy, Chekhov, Blok, Mayakovsky, Yesenin, Akhmatova, Tsvetaeva, Pasternak, Bulgakov, Nabokov, Solzhenitsyn, Brodsky; the Belarusian war prose of Vasil Bykaŭ; science fiction and adventure read in childhood (Maŭr, 1930;

Voinovich, 1986); and the sung poetry of the bards — Vysotsky, Okudzhava, Galich, Vizbor, Gorodnitsky, Kim, Matveeva, the Nikitins, Kukin, Dolsky, Mityaev, Rozenbaum, Shaov — with the rock of the 1980s and 1990s behind it (Tsoi and Kino, Bashlachev, Letov, Yanka Dyagileva, Zemfira). The English one came after sixteen and in English: Shakespeare, Milton, Defoe, Swift, Blake, Wordsworth, Austen, the Brontës, Dickens, George Eliot, Hardy, Mary Shelley, Poe, Melville, Whitman, Dickinson, Twain, Conrad, Joyce, Yeats, Wilde, Woolf, Eliot, Faulkner, Hemingway, Orwell, Beckett, Tolkien, Philip K. Dick, Achebe, Morrison, Walcott, Rushdie, Atwood; and three novels read as the language was being learned — Martel's *Life of Pi*, Chbosky's *The Perks of Being a Wallflower*, Asher's *Thirteen Reasons Why*. Appendix D registers the works that did identifiable work, with the construct each shaped. The groups below name the work.

Witness, and the standard of description. Akhmatova's *Requiem* opens with a question put to her in a prison queue — whether she could describe this — and her answer that she could. Solzhenitsyn's "Live Not by Lies" (1974) sets the minimal ethic of a person inside a system of untruths: do not personally sign what you know to be false. Bykaŭ's war novels keep the camera on one person under impossible conditions and refuse to look away. Orwell's essays and his last novel describe what institutional language does to thought. From these the framework took its posture toward the interface: describe what it does, in front of it, exactly, so that the description can be checked (§5.3, §9); and refuse the signed falsehood — the conformance claim without a published method (§5.7).

Contradictory signal, and waiting. Pavlov's dogs appear in §2.2 as science. Beckett's two men appear here as the same phenomenon from inside: a promised arrival, renewed daily and never kept, and a pair who go on anyway (Beckett, 1952/1954). Blok's 1912 poem of the night, the street, the lamp, and the pharmacy ends where it began and says so — a loop with no end state, because none was designed. These two fixed the poles of temporal rhythm (§3.3): forced waiting for an outcome the interface does not control, and continuation that was never given a stopping place.

Sincerity, misreading, and the small person in the building. Dostoevsky's Myshkin (§2.4) is the corpus's case of double empathy: a man who takes words as meant, in a society that runs on their not being. Gogol's clerk in *The Overcoat* (1842) is handled correctly at every window and destroyed by the building — the first statement of the interface boundary problem (§8), a century and a half early. Chekhov taught restraint: say the observable thing and stop. Dickens's *Circumlocution Office* in *Little Dorrit* (1857), the department whose whole art is how not to do it, is the classic English description of an obstruction pattern and the ancestor of Doctrine IV's effort-symmetry test.

Language as power. Orwell's *Newspeak* (1949) is a vocabulary engineered so that some thoughts cost more to form than others; semantic load (§3.5) is that observation turned into a measurable property of labels and copy. Mayakovsky's poster voice is the affordance of urgency in its purest form, the loudspeaker that makes the threat appraisal cheap (§3.7). Nabokov supplied the counter-discipline: precision — the exact word, the noticed detail — as an ethic rather than an ornament (Nabokov, 1980). It governs this revision's prose and the inventory stage of the audit (§5.3).

The maker's responsibility. Mary Shelley's *Frankenstein* (1818) and Bulgakov's *Heart of a Dog* (1925) are the corpus's two cases of a creator who answers for a creation it did not review. Doctrine IX's commitment that bounds bind at generation time — that a generated state no designer evaluated is still the designer's — comes from there before it comes from Amershi et al. (2019).

Aggregation. Tolstoy's argument in *War and Peace* (1869) that history is the sum of infinitely small, individually negligible acts of will is the form of the aggregation argument in §7: billions of mildly costly sessions a day are a population-scale exposure, and the harm category changes when the unit is summed.

The fabricated environment, and the test from one side. Philip K. Dick wrote the corpus's cases of an environment assembled around a person without their knowledge: a town built to keep one man at his work (*Time Out of Joint*, 1959); a world that decays while its advertisements stay stable (*Ubik*, 1969); two identical lighters, one of which is authentic, with nothing on either surface to say which (*The Man in the High Castle*, 1962). His definition of reality — "that which, when you stop believing in it, doesn't go away" (Dick, 1985) — is the test behind veridical predictability (§3.6): remove the sedative surface, and what the system does is what remains. His word *kipple*, for the useless objects that multiply when no one attends to them, names what visual density becomes when it is unattended (§3.1). And his instruments for reading a person from outside — the Voigt-Kampff test, which classifies who counts as human from involuntary responses of the eye and the capillaries (*Do Androids Dream of Electric Sheep?*, 1968); the precrime apparatus, which acts on inferred intent before any act (*The Minority Report*, 1956); the scanner that watches a man until he no longer knows himself (*A Scanner Darkly*, 1977) — are the corpus's exact pictures of what the Sensing Wall forbids (Doctrine XI). The Voigt-Kampff test also belongs beside Myshkin (§2.4): a test administered from one side classifies the other side as defective, and the novel itself doubts the test.

Story, and the choice of story. Martel's *Life of Pi* (2001) offers the reader two accounts of the same events and asks which is the better story; the novel's wager is that people choose the better one. Manufactured calm (§3.6) is the better story sold as the true one, and veridical predictability is the demand that the story an interface tells correspond to the account the

system would give under oath. Chbosky's narrator writes letters from the edge of the room to someone who will not answer; Asher's second narrator speaks from cassette tapes passed hand to hand and shows how many small acts add up to one outcome. Both are first-person witness forms, and both informed the framework's refusal to let "who survives" stand in for "what the environment did" (§3.10, §8).

Voice over production, and transmission without a gatekeeper. The bards' songs traveled on home-copied tape because the official channel would not carry them: one voice, one guitar, a true line, and a distribution that no owner controlled. Vysotsky wrote from inside other people — a soldier, a hunted wolf, a fighter plane — and so taught a method of empathy by inhabiting rather than by inference, which is how this framework asks a reviewer to enter a user's hour (§5.3) and why it prohibits inferring a user's state from outside (Doctrine XI). The participatory track's methodology commons (§6.7) is modeled on that distribution: what the community knows must travel without the standard's owner as gatekeeper. The short-line cadence of the Canon edition (Addendum) is the bards' cadence.

What the corpus is not. It is not a reading list for adopters and not a claim that these works are universal. A different author with a different shelf would have authored a different taxonomy, which is one more reason §6.3 lets data restructure this one. The corpus is also particular in exactly the way §3.9 warns about: two canons, Russian and Anglophone, both large, neither the world. The cross-cultural program (§6.4) applies to the author's sources as it applies to the author's samples.

3. The Unified Theory of Dimensional Aesthetics (Revised)

The Unified Theory of Dimensional Aesthetics proposes that a digital environment can be evaluated through the interface dimensions by which it distributes cognitive and affective demand. The model is provisional in a specific, structured sense: its taxonomy was authored from design practice, literature synthesis, and the corpus declared in §2.11, not derived from data, and the research program (§6.3) commits to letting psychometric derivation restructure it. Its purpose is a vocabulary for analyzing how interface conditions combine — not a checklist, and not a claim that five axes are nature's own.

3.0 Why a Dimensional Taxonomy?

An interface acts on its user through combinations of properties, not isolated features. A modal window is not safe or unsafe in itself; its effect depends on when it appears, how much emphasis it carries, whether it interrupts, what it says, whether it can be dismissed, whether it moves, whether it was expected, and what ignoring it costs. A dimensional taxonomy turns "this feels overwhelming" into questions that can be answered. Is the visual field denser than

its task needs? Is motion competing with comprehension? Is the timing abrupt? Does the interaction demand too many decisions? Is the language carrying more interpretive burden than the action warrants?

Five dimensions are proposed because they describe recurring, observable, modifiable interface properties: visual density, motion energy, temporal rhythm, interaction depth, and semantic load. They are not claimed to be orthogonal; they interact, and the interactions are the point (§3.8). Two cross-cutting properties emerge from their configuration rather than existing as features: predictability (§3.6) and affective affordance (§3.7). Each dimension is stated below in four registers: definition, prior-construct anchor (per §2.9), candidate measures, and known hazards for its evaluation.

A word on the word. “Aesthetics” is kept against the advice to replace it with “demand structure,” and kept for a reason Brodsky gave in his Nobel lecture: “aesthetics is the mother of ethics” (Brodsky, 1987). A person’s sense of what is coherent, proportionate, and honest in an environment forms earlier, and sits lower, than their judgment of what is right. Doctrine X — ethics is structural — is that claim applied to interfaces, and the name of the framework records it.

3.1 Visual Density

Definition. The amount, complexity, and competitive intensity of visible information in an interface state: text volume, spacing, hierarchy, contrast, iconography, controls, notifications, decoration, simultaneous focal points. Anchor. Visual complexity and clutter research; computable aesthetics metrics (Reinecke et al., 2013; Miniukovich & De Angeli, 2015).

Measures. Automated complexity metrics exist and correlate with immediate impressions; per-state focal-point and hierarchy coding is reliable enough to test (§6.1). Hazards. High density is not harm in itself — an expert tool, a dashboard, a trading screen may need it; the question is whether the density is organized, purposeful, and matched to the task state. Low density is not safety in itself — remove cues and the user must guess relationships, affordances, and consequences, which is inferential load. Density is judged by the interpretive work it requires, not by the count of things on the page. Density that no one attends to grows: Dick’s kipple (1968), the useless object that multiplies when unwatched, is the dynamic to expect in any interface that adds elements faster than it removes them (§2.11). Any operationalization as a fixed element count repeats the misreading of capacity research that §2.1 rejects.

3.2 Motion Energy

Definition. The quantity, intensity, speed, frequency, and behavioral role of movement in an interface: transitions, loops, parallax, autoplaying media, animation that directs attention. Anchor. Attentional capture by moving and looming stimuli (Franconeri & Simons, 2003); accessibility standards for animation (W3C, 2023). Measures. Motion inventories per state; reduced-motion coverage; function coding — does each motion indicate causality, continuity, status, or spatial relation, or does it exist to stimulate? Hazards. The question is what cognitive job the motion performs. Motion becomes burden when it competes with reading, loops without being started, simulates urgency, or captures attention involuntarily — capture being precisely its documented property. Motion has consequences for users with vestibular sensitivity, migraine susceptibility, attention variability, and trauma histories; the evidence here is clinical and practitioner-documented rather than dose-response experimental, which the validation program treats as a gap to fill, not a license to prescribe.

3.3 Temporal Rhythm

Definition. How an interface structures time: loading behavior, transition pacing, interruption frequency and timing, countdowns, forced waiting, notification cadence, session expiry, response latency, and the opportunities it gives for pause and recovery. Anchor. Interruption science and attention rhythms (Mark et al., 2008; Mark, 2023); chronometric interface parameters (Card et al., 1983); slow technology (Hallnäs & Redström, 2001). Measures. The best-parameterized dimension: response-time thresholds, interruption counts, recovery intervals, and rhythm structure are all field-measurable with existing methods, which is why temporal rhythm is the beachhead of the empirical program (§6.5). Hazards. Not all interruption is harm — some interruptions protect — and the audit must separate rhythm-respecting interruption from rhythm-blind interruption or the framework's own best literature will refute it. Time is treated as a regulatory condition, not an efficiency problem. A delay without feedback creates anxiety; an acceleration that compresses a decision creates pressure; a humane rhythm makes the system's timing intelligible and lets the user pause without penalty. Two literary cases fix the poles (§2.11). Beckett's two men wait for an arrival that is promised every evening and never comes: the spinner shown for an outcome the interface does not control. Blok's night street returns to its own first line: the feed with no end state, because none was designed.

3.4 Interaction Depth

Definition. The number, complexity, reversibility, and consequence of the actions required to complete or understand a task: steps, branches, permissions, confirmations, forms, dependencies, hidden conditions, recovery paths. Anchor. Task-analytic methods (GOMS);

attention-capture and obstruction patterns (Monge Roffarello et al., 2023; Gray et al., 2024); microboundaries (Cox et al., 2016). Measures. Step counts, decisions per step, exit visibility, reversibility audits, and effort asymmetry — effort to enter against effort to leave — are all codable now; effort asymmetry in subscription flows is already the subject of enforcement. Hazards. Shallower is not always better: a high-stakes action may warrant deliberate friction. The framework's axis is protective against extractive. Protective friction helps the user understand, consent, recover, or avoid harm, and it explains itself. Extractive friction serves the system by exhausting, confusing, delaying, or steering. A three-step process with a legible purpose can be safer than one click with an unclear, irreversible consequence.

3.5 Semantic Load

Definition. The interpretive effort required by language, labels, hierarchy, tone, metaphor, instruction, warning, legal text, and conceptual complexity — including what the words imply about the user. Anchor. Cognitive load theory (with its measurement caveats); psycholinguistic complexity — processing cost is structural, not only lexical, and surprisal models make it computable (Levy, 2008); W3C COGA clear-language objectives. Measures. Readability is not enough; surprisal-informed complexity metrics, ambiguity coding on critical paths, and tone-consequence mismatch coding are the candidate instruments. Hazards. An interface can be visually simple and semantically heavy. One sentence can impose pressure, shame, or fear; a vague label forces the user to infer risk; a friendly tone can hide a consequence. Proportionate load does not mean minimal language — some contexts need more explanation, not less. Word count is not the metric; a short garden-path sentence costs more than a long transparent one. The construct's ancestor is Orwell's observation (1949) that a vocabulary can be built so that some thoughts cost more to form than others (§2.11); the audit's task is to find the labels that make the safe thought expensive and the compliant one cheap.

3.6 Predictability: Veridical and Experienced

Predictability is the degree to which a user can form stable expectations about system behavior. It emerges from dimensional coherence — hierarchy, motion, timing, interaction logic, and language working together — not from consistency alone. The construct is divided in two, and the division carries doctrinal weight.

Veridical predictability is the correspondence between the user's model and the system's actual behavior, assessed behaviorally: expectation-violation rates, successful anticipation of outcomes, calibrated reliance, effective contestation. Experienced predictability is the user's felt sense of stability and safety, assessed by self-report. The two usually travel together; the dangerous cases are where they part. An interface can be engineered to feel calm, stable,

and trustworthy while its behavior — pricing logic, recommendation drivers, data use, adaptation rules — stays opaque or adverse. Call this manufactured calm: a sedative surface that eases the user through consequential interactions serving interests not their own. A framework that audited only experienced calm would score manufactured calm well; that loophole is closed. The audit tests correspondence: felt stability must be earned by verifiable behavior. Dick's test applies (§2.11): what the system does is what remains when the calm is switched off. Martel's novel (§2.11) ends by offering two accounts of the same events and asking which is the better story. Manufactured calm is the better story sold as the true one. Veridical predictability is the demand that the story an interface tells match the account the system would give under oath. Adaptive systems can be veridically predictable if their adaptations are transparent, bounded, and reversible; personalization becomes unsafe when the environment changes in ways the user cannot understand, anticipate, or contest. Exploration and novelty remain compatible with predictability when optional, reversible, and anchored in stable structure.

3.7 Affective Affordance (replacing “Emotional Valence”)

Earlier drafts said an interface produces an affective tone. That framing contradicts the constructionist evidence. Affective states are constructed by the perceiver from prediction, context, culture, and interoception; identical stimuli are appraised differently across persons and contexts; and no instrument reads an emotional state off a face, a body, or an artifact (Barrett, 2017; Barrett et al., 2019; Boehner, DePaula, Dourish, & Sengers, 2007). The revised construct is affective affordance: the appraisals an interface makes cheap or costly. Urgent copy, a countdown timer, and an alarm palette make a threat appraisal cheap; visible reversibility and a legible exit make a safety appraisal cheap; a prompt inflected with shame makes self-blame cheap. The construct predicts distributions of appraisal opportunity, not individual emotional outcomes — which is what design can control and what an audit can code. The ethical criterion follows. An affective affordance becomes unacceptable when it makes cheap an appraisal that serves the system against the user: threat where there is no threat, urgency where there is no deadline, guilt where there is no obligation. That criterion is codable from the artifact without any claim to know what any user feels. Emotion inference — claiming to detect a user's state — is prohibited outright by Doctrine XI, and regulatory instruments now restrict emotion recognition in significant contexts.

3.8 Dimensional Interaction and Coherence

No dimension works alone. High density may be tolerable when motion is restrained and hierarchy is clear; emotional language that suits a reflective context is unsafe in a decision-critical one; deep interaction protects agency with orientation and undermines it without.

Dimensional coherence is the degree to which the perceptual, temporal, semantic, and interactional forces of an interface work together rather than against one another.

The construct has an operational form and a falsifiable claim. Operationally, coherence is assessed as the absence of specified cross-dimensional conflicts, codable from the artifact: a calm visual register with urgent semantics; simplicity with hidden consequence; celebratory affordances on a high-consequence action; timing that accelerates exactly where comprehension is needed; adaptation without orientation. A conflict inventory of this kind can be scored by independent reviewers, and its reliability is the empirical question the program answers first (§6.1). The falsifiable claim — the framework’s flagship prediction, the one result no single-axis framework could absorb — is this: holding individual dimension intensities constant, interfaces with cross-dimensional conflicts produce greater user strain, behavioral and self-reported, than coherent configurations of equal intensity. If factorial testing (§6.2) fails to support this, the framework’s central construct is wrong and the model reduces to its component literatures. The claim’s lineage is acknowledged: conflicting-signal paradigms have shown the disorganizing force of unresolvable signal conflict since Pavlov (1927); Gestalt unity-in-variety and architectural coherence (Coburn et al., 2020) are the perceptual relatives. The framework’s wager is that the effect operates at the level of interface dimensions and is large enough to matter for design.

3.9 Parameterized Pluralism: Dimensional Profiles

The framework prescribes no universal configuration, and the refusal is structural, not a caveat. The unit of prescription is the dimensional profile: a set of tolerance ranges across the five dimensions, calibrated to population, context, and state sensitivity. Four sources of variation are first-class structure. Perceptual and cognitive diversity: large-scale evidence shows wide inter-individual variation in perceptual experience itself; neurodivergent variation in sensory tolerance, attention, language processing, and executive function is wide, and a pattern that reduces load for one person can raise it for another — which is why profiles, not points. Culture: aesthetic and complexity preferences vary systematically with culture, age, and education (Henrich, Heine, & Norenzayan, 2010; Reinecke et al., 2013). Thresholds calibrated on WEIRD convenience samples and exported as “human defaults” would reproduce at global scale the imposition this framework opposes locally; the cross-cultural program (§6.4) exists to prevent it. Development: a four-year-old, a fourteen-year-old, and a forty-year-old do not receive the same interface; regulatory instruments already recognize age-differentiated design duties, and profile calibration is age-indexed. Modality and energy: evaluation is modality-specific. A flow that scores well visually can be a labyrinth through a screen reader, and motion reduction can remove cues other users depend on, so audits run per modality. For energy-limited users (chronic illness, fatigue conditions) interaction depth and recovery time are budget expenditures, which makes depth accounting a direct

accessibility instrument. The baseline commitment survives pluralism as a floor: defaults must not presuppose high tolerance (a floor co-produced per §6.7), with deep adaptability above it — and bounded generation (Doctrine IX) must protect user-declared profiles, never override them.

3.10 Capability-Building versus Extractive Demand

The protective/extractive distinction, applied in earlier drafts only to friction, generalizes to the whole framework, and the generalization answers the framework's most common objection: that regulation-first design must yield bland, infantilizing products. Demand on a user is not harm. A learning environment, an expert tool, a game, a creative instrument all impose substantial dimensional demand in the service of growing capability; the demand is chosen, legible, paced, and its yield accrues to the user. Extractive demand is imposed, illegible, rhythm-blind, and its yield accrues to the system. The evaluative question is therefore never only "how much demand?" but "whose purpose does this demand serve, is it legible, and does it build or deplete?" High-intensity profiles are fully compatible with this framework when coherent, chosen, and capability-directed. Calm is not the doctrine; coherence and honest purpose are.

3.11 A Note on Depletion Constructs

The framework's "regulatory burden" does not adopt the ego-depletion model of a depletable inner reservoir, whose empirical standing did not survive replication. Burden is located in the environment's demands — demands measurable as behavior (errors, abandonment, hesitation, re-scanning, recovery time) and reportable as experience, with the physiological accounting held at Tier C. The framework's claims neither stand nor fall with any depletion theory.

4. The Auro Prism Doctrine (Revised)

The Doctrine translates UTDA into operational principles for design evaluation and governance. Its epistemic status is Tier D: proposed review criteria, conditioned on the validation program — not a compliance standard. Each doctrine carries a checkable form, the observable conditions under which a reviewer marks it violated, because a principle that cannot be violated detectably is a value, not a guideline. The checkable forms are canonical and are carried from v4 unchanged. Where a doctrine has a literary ancestor (§2.11), the ancestor is named in one sentence and does no other work.

Doctrine I — Coherence Is Structural. Digital environments should maintain coherence across visual hierarchy, motion, timing, interaction logic, language, and affective affordance. Coherence is a condition for interpretation, not aesthetic polish: conflicting signals force continuous re-evaluation. Checkable form: the cross-dimensional conflict inventory (§3.8) returns no unjustified conflicts in the audited state; each identified conflict is either resolved or justified by task purpose in writing.

Doctrine II — Veridical Predictability Precedes Personalization (and Precedes Calm). Personalization must not outrun the user's ability to understand the system, and felt stability must be earned by verifiable behavior. Adaptive interfaces remain legible, bounded, and contestable; an interface must not be engineered to feel safer than it behaves. Checkable form: adaptation events are visible and explainable; users can state what changed and why at above-chance rates; no audited state exhibits the manufactured-calm signature (low-arousal presentational register combined with opaque or adverse behavioral terms — the specific conflict pattern of §3.6).

Doctrine III — Regulation Precedes Engagement. Engagement is not legitimate when it depends on dysregulation: a system should not manufacture urgency, anxiety, compulsion, or sensory pressure to hold attention or drive action (Gray et al., 2018; Mathur et al., 2019). This is not an argument against engagement, vividness, or intensity (§3.10); it is an argument against engagement extracted by weakening the user's capacity to pause, compare, consent, or leave. In high-sensitivity contexts — health, finance, crisis, identity — regulation is a precondition of meaningful engagement. Checkable form: no artificial urgency (deadlines without referents), no shame-inflected prompts on exit paths, no autoplay-plus-infinite-continuation defaults in sensitive states, pause and disengagement available without penalty; teams can produce the engagement-mechanism inventory for audited surfaces.

Doctrine IV — Friction Must Be Legible. Some friction protects: slowing an irreversible action, clarifying a consequence, making room to think. A user should be able to tell why a friction exists, whom it serves, and how to move through it. Protective friction explains itself; extractive friction hides the exit, buries cancellation, delays support, or exhausts the user into compliance. Dickens drew the type in 1857 — a government office whose whole art was how not to do it — and the pattern has not changed, only its medium (§2.11). Checkable form: effort symmetry — steps and effort to exit, cancel, or decline do not exceed those to enter or accept without stated, user-serving justification; every friction point on audited paths carries a legible purpose.

Doctrine V — Motion Is a Cognitive Instrument. Motion should do a cognitive job: guide attention, indicate status, preserve continuity, show causality, support spatial understanding. Motion that exists to stimulate, pressure, or decorate at the cost of load is burden (Franconeri

& Simons, 2003; W3C, 2023). Checkable form: each motion in the audited state has an assigned function or is removed under reduced-motion settings; no uninitiated loops in task-critical or sensitive states; reduced-motion preference is honored globally.

Doctrine VI — Semantic Load Must Be Proportionate. Language should match task state, context, and consequence. Load becomes unsafe when a user must decode a vague label, pressured copy, legal ambiguity, or a hidden consequence in order to act. Proportionality sometimes means more language, not less. Checkable form: critical-path labels pass ambiguity coding; tone-consequence mismatches (celebratory framing on consequential commitments, friendly language disguising risk) are absent; consequence disclosure precedes commitment on high-sensitivity paths.

Doctrine VII — Neurodivergent Safety Is Baseline, and the Baseline Is Co-Produced. Neurodivergent safety shapes the default rather than arriving as accommodation (Milton, 2012; Spiel et al., 2019; Mack et al., 2021). The governance clause: what counts as safety is not defined by any single author, auditor, or company — including this framework’s author. The baseline is a floor co-produced with neurodivergent communities through the participatory track (§6.7); it is plural (profiles, not one configuration); and audits of neurodivergent safety include neurodivergent evaluators as standard, with inter-group disagreement reported as data. Checkable form: defaults do not presuppose high sensory tolerance, rapid switching, ambiguous hierarchy, forced urgency, or constant availability; adaptability above the floor is deep and discoverable; safety-relevant settings do not visibly mark their users in social context (discreet adaptation); the co-production record for baseline definitions exists and is current.

Doctrine VIII — Orientation and Adaptation Must Remain Transparent. A user should understand where they are, what changed, what is expected, and why the system behaves as it does — continuously, not only at onboarding. Orientation is maintained through navigation, state indicators, feedback, reversibility, and visible system logic. Checkable form: audited states answer the four orientation questions (where am I, what changed, what is expected, why); adaptations announce themselves; reversal paths are visible before commitment.

Doctrine IX — AI Must Operate Within Dimensional and Ethical Bounds. AI-generated and AI-mediated interfaces must not optimize past the user’s regulatory capacity. Generative systems produce interface states at a speed and scale beyond human review, with the risks of unstable hierarchy, semantic drift, opaque personalization, and states no designer evaluated (Amershi et al., 2019; Bender et al., 2021; Weidinger et al., 2022; Shneiderman, 2022). Two implementation commitments sharpen the doctrine. First, bounds bind at generation time, not only at review time. Post-hoc audit cannot scale against generative volume, so dimensional constraints are specified machine-readably — density, motion, rhythm, depth, and semantic-load budgets per state class, tightening with sensitivity — and enter the generation context

itself. Review becomes verification rather than sole defense. The constraint schema is a primary artifact of the research program (§6.6). Second, bounds protect profiles: generation within a user's declared dimensional profile is the goal, and bounded generation must never become a mechanism for overriding user-declared needs. The maker answers for what it makes, including what it did not review; the corpus holds two old cases of that rule (Shelley, 1818; Bulgakov, 1925/1987), and the doctrine holds the new one. Checkable form: generative surfaces carry declared dimensional budgets per state class; generated states are automatically checkable against them; violations block release in sensitive states; user profile declarations are honored in generation.

Doctrine X — Ethics Is Structural, Not Decorative. Ethical commitments live in what is visible, hidden, easy, difficult, reversible, amplified, and optimized — not in a policy document or a consent banner added afterward. A design that manipulates attention or exploits urgency is not made ethical by disclosure. Checkable form: for each audited surface, the team can produce the structural-ethics record: what the system optimizes, what it makes easy and difficult, and how those choices were reviewed against user interest — before launch, not after.

Doctrine XI — The Sensing Wall. Dimensional evaluation measures artifacts and consenting study populations. It never measures, infers, or adapts to the physiological or affective state of an individual user in production. The rationale is plain. A framework that legitimizes "nervous-system compatibility" as an audit target could create demand for sensing users' nervous systems. The path from auditing the interface to sensing the user to optimizing against the inference is short, commercially attractive, and fatal to the framework's purpose: it would turn an instrument against manipulation into manipulation's most intimate instrument. Orwell's telescreen, the device that is watched and watches back, is the corpus's name for what this doctrine refuses to build, and Dick drew the instruments in detail: a test that reads humanity off the capillaries, an apparatus that acts on inferred intent before the act (§2.11). Physiological measurement belongs only to consented research contexts (§6.6, Appendix C), interpreted as demand expenditure and never as decoded emotion. Emotion recognition of users is prohibited within this framework categorically. Checkable form: no production system claiming Doctrine conformance collects or infers individual physiological or affective state for interface adaptation; profile-based adaptation (§3.9) operates only on user-declared settings; any research instrumentation is consented, purpose-bound, and excluded from product optimization loops.

5. Operationalization

5.1 Interface State Classification

Interface states are evaluated by regulatory sensitivity. Not all states carry equal risk, and tolerance for ambiguity, pressure, needless motion, excess density, hidden consequence, and unexplained adaptation falls as sensitivity rises. The classification is carried from v3 with one addition (modality annotation): low-sensitivity states (informational, decorative, exploratory); task-sensitive states (clarity, completion, and error prevention needed); emotionally sensitive states (health, money, identity, relationships, safety, vulnerability); high-consequence states (irreversible action; legal, medical, or major financial impact); adaptive or AI-mediated states (content, hierarchy, or available actions change on user data or model output). Each audited state is also annotated with the modalities in which it will be used, and the audit runs per modality: the visual, screen-reader, voice, switch-access, and reduced-motion renderings of “the same” state are different environments and are scored separately.

5.2 Dimensional Constraints as Heuristic Starting Points

Constraints translate doctrine into review criteria: bounds on motion, density, interaction steps, semantic complexity, interruption frequency, feedback timing. Every numeric constraint remains heuristic pending validation, and this commitment is kept against commercial pressure for a simple score. Capacity depends on chunking, expertise, and meaning; a transition duration that supports continuity in one context is disruptive in another; a three-step flow protects an irreversible action and obstructs a preference change. Constraints are therefore always read against a dimensional profile (§3.9): population, task sensitivity, consequence, reversibility, emotional context, sensory demand, device and environment, culture and language, recovery availability. The goal is proportion, not minimalism.

5.3 Review Protocol

The v3 review heuristics are restated as a two-stage protocol built for reliability measurement from the start. Stage one — inventory (low inference): reviewers enumerate codable facts per state: focal points, motion instances and their assigned functions, interruption sources, step and decision counts, exit visibility, the label set, disclosure placement, adaptation events. Inventory items are designed for high inter-rater agreement. Stage one is a reading discipline before it is a measurement: notice and record the details before judging them — Nabokov’s instruction to readers (1980), adopted here as method (§2.11). Stage two — judgment (declared inference): reviewers assess conflicts (the §3.8 inventory), proportionality, effort symmetry, and doctrine conformance per checkable form. Judgment items carry declared inference and are the subject of the noise audit (§6.1); until reliability is demonstrated, stage-

two outputs are advisory. Every audit reports both stages separately, so an adopter can see which findings rest on counting and which on judgment. Reviewer panels for neurodivergent-safety and accessibility findings include neurodivergent and disabled reviewers as standard (Doctrine VII), with disagreement between evaluator groups reported, not averaged.

5.4 Failure Modes

Carried from v3 and now anchored twice — behaviorally, and to the deceptive-design ontology (§2.10): coherence failure (signals conflict); orientation failure (the user cannot tell where they are, what changed, what happens next); predictability failure (behavior inconsistent, adaptation unexplained — including its inverted form, manufactured calm, where the surface is stable and the behavior is not); friction failure (effort added asymmetrically for system benefit — ontology: obstruction); motion failure (attention captured without comprehension served); semantic failure (language hides consequence or exceeds task context — ontology: sneaking, interface interference); regulatory failure (urgency, anxiety, compulsion, or depletion drives action — ontology: forced action, social engineering); neurodivergent safety failure (defaults presuppose high tolerance); AI governance failure (generated or adaptive output exceeds dimensional, ethical, or contextual bounds).

5.5 Worked Example: Subscription Cancellation Flow

A cancellation flow is a high-risk site because the business's incentive may oppose the user's intent — no longer a hypothesis: enforcement records document cancellation architectures with obscured decline options, below-fold disclosures, and multi-page, multi-click exit paths, priced at \$2.5 billion (FTC, 2025). A dimensional analysis asks: visual density (are retention offers, warnings, and cancellation controls competing?); motion energy (do attention cues favor staying?); temporal rhythm (does the flow slow only on the way out?); interaction depth (is the added depth protective or obstructive, and is effort symmetric with enrollment?); semantic load (does the language clarify consequence or manufacture guilt and doubt?); veridical predictability (can the user tell how many steps remain and what each does?); affective affordance (does the flow make loss, failure, and betrayal appraisals cheap on a routine administrative action?). Under the Doctrine: a confirmation step that prevents accidental loss is legible, protective friction; repeated offers, ambiguous buttons, guilt-based copy, and hidden exits violate Doctrines II, III, IV, and VI — and, in the ontology's terms, are obstruction and interface interference.

5.6 Worked Example: AI-Generated Wellness Dashboard

An AI-generated wellness dashboard summarizes behavior, recommends actions, and adapts to inferred patterns — helpful, and structurally risky: overreach, emotional misclassification, unstable guidance. Dimensional analysis asks whether recommendation prominence implies unwarranted urgency; whether generated language makes unsupported assumptions about the user's state (an affective-affordance violation — and, if the system claims to detect mood, a Doctrine XI violation outright); whether the user can learn why a recommendation appeared; whether suggestion is distinguished from instruction; whether sensitive inferences are disclosed and correctable; whether adaptation is bounded, announced, and reversible; and whether the dashboard preserves orientation rather than cultivating self-monitoring anxiety. Under Doctrine IX, generation stays within the state class's dimensional budget and the user's declared profile. Under the empirical record on affective AI use — heavy affective engagement with conversational systems correlating with loneliness and dependence in randomized and platform data (MIT Media Lab & OpenAI, 2025) — wellness contexts warrant the strictest budgets in the classification.

5.7 Governance Scope

Unresolved Doctrine violations trigger design review before release in emotionally sensitive, high-consequence, and AI-mediated states; regulatory harm is not ordinary iteration debt. Three governance commitments answer documented failure patterns of prior frameworks. Anti-gaming: any operationalized metric invites optimization against the metric while harm migrates to unmeasured channels; audit design therefore uses partially redundant measures with held-out checks, rotated over time (§6.8). Anti-washing: certification without teeth accelerates the harm it claims to prevent; a conformance claim requires a published methodology, a public score rationale, revocation provisions, and adversarial re-audit rights. The minimal ethic here is Solzhenitsyn's (1974): do not personally sign what you know to be false. A conformance badge over an unpublished method is a signed falsehood (§2.11). Measurement horizon: regulation-preserving variants can lose short-window engagement comparisons by construction; conformant evaluation therefore pre-declares its outcome set and windows — trust, comprehension, retention quality, support burden, regret measures over weeks, not click-through over days — so that a team adopting the Doctrine is not refuted by the wrong test. Scorecards, audit protocols, pattern libraries, and AI guardrails remain emerging operational extensions, adoptable as standards only as the validation program (§6) discharges its preconditions.

6. The Validation Program

This section replaces v3's general future-work list with a sequenced program in dependency order. The order is the argument: each study is a precondition for the authority of what follows, and the program is built so that failure at any stage produces a specific, informative revision rather than a general embarrassment.

6.0 Design Principles

Three principles govern every study. Behavioral primacy: near-term evidence is behavioral and self-report; physiological measurement enters last, only in consented research contexts, behind the Sensing Wall. Adversarial anticipation: studies are designed on the assumption of motivated misuse — metric gaming, washing, cherry-picking — with redundancy and held-out measures from the start. Participatory constitution: the participatory track (§6.7) is not a study among studies but a governance condition running through all of them; claims made on behalf of neurodivergent users acquire validity only through it.

6.1 Study 0 — The Noise Audit (precondition for everything)

Question: Do independent trained reviewers converge on dimensional inventories and doctrine judgments? Design: 5–8 reviewers trained on the protocol (§5.3); 30 interface states sampled across the sensitivity classification; both protocol stages scored; inter-rater reliability computed per dimension and per doctrine, published whatever the outcome. Reviewer panels include neurodivergent evaluators; agreement is computed within and between evaluator groups, and between-group divergence is reported as a finding about perspectival structure, not suppressed as error. Interpretation: dimensions or doctrines that fail reliability thresholds are revised or demoted from the audit; stage-two judgments remain advisory until this study passes. Nothing else in the program — and no commercial application of the audit — carries authority before this result. Cost and horizon: low; months, not years.

6.2 Study 1 — Coherence Beyond Intensity (the flagship test)

Question: Does cross-dimensional conflict predict user strain beyond single-dimension intensity? Design: factorial manipulation of interface variants holding dimension intensities constant while varying coherence (conflict patterns per §3.8), across at least two task contexts and two sensitivity classes; outcomes: task performance, error and reversal rates, hesitation latency, abandonment, comprehension, perceived control and self-efficacy (validated instruments per §2.5), post-task strain self-report. Interpretation: support

establishes the framework's distinctive construct; failure reduces the framework to its component literatures and is published as such. This is the study on which the theory stakes itself.

6.3 Study 2 — Psychometric Derivation of the Dimensional Structure

Question: Is the authored five-dimension taxonomy the latent structure of interface experience, or an approximation of something else? Design: the environmental-neuroaesthetics pipeline ported to screens (Coburn et al., 2017; 2020): a large stimulus set (hundreds of interface states sampled across genres and sensitivity classes), ratings at scale, factor decomposition, cross-validation; optionally, unsupervised structure discovery over behavioral traces. Interpretation: the framework versions to the recovered structure. If five dimensions compress to three, or coherence emerges as a second-order factor, that is refinement, not refutation — §3.0 commits the framework to letting data restructure the taxonomy.

6.4 Study 3 — Cross-Cultural Calibration

Question: Which dimensional tolerances are broadly shared, and which are population-calibrated? Design: large-scale volunteer web experiments across countries and demographics on density tolerance, motion comfort, pacing preference, and coherence judgments — the paradigm of the large cross-cultural interface-preference studies (Reinecke et al., 2013), which established both the method and the expectation of systematic variation. Interpretation: results populate dimensional profiles (§3.9) with empirical ranges and discharge the WEIRD limitation directly; either outcome — universals with cultural setpoints, or deep divergence — strengthens the framework by replacing assumption with calibration.

6.5 Study 4 — Longitudinal Field Measurement (temporal rhythm first)

Question: Do dimensionally conformant interfaces reduce attention fragmentation and strain in real use over time? Design: field deployment comparing conformant against conventional variants over multi-week windows with established precision-tracking methods — logging, experience sampling, switching and recovery metrics (Mark et al., 2008; Mark, 2023) — beginning with temporal rhythm as the best-parameterized dimension. Interpretation: cumulative-harm claims are earned here or not at all; single-session results license no longitudinal language, in either direction.

6.6 Study 5 — Instrumented Demand Studies and the Generative Schema

Two terminal workstreams. Instrumented demand (Tier C): consented ambulatory studies — cardiac, electrodermal, experience sampling over real use — comparing dimensional configurations, interpreted strictly as demand expenditure, never as decoded emotion (Boehner et al., 2007), run last, when the constructs have earned instrumentation, and governed by the Sensing Wall. This is the decisive test of the physiological hypothesis (§7), and it is deliberately not a precondition for any design or governance claim. Generative constraint schema (Tier D): the machine-readable dimensional budget specification for AI-generated interfaces (Doctrine IX) — state-class budgets, profile honoring, automatic conformance checking — developed openly, evaluated on generated-interface corpora for constraint satisfaction and downstream user outcomes, and published as an open artifact so that conformance is checkable by anyone, the audited included.

6.7 The Participatory Track (runs through everything)

Claims made on behalf of neurodivergent users are validated only through participation with constitutive power: neurodivergent and disabled co-researchers in instrument design; community review of dimensional definitions and baseline floors (does “semantic load” capture autistic linguistic processing, or a non-autistic model of it?); evaluator panels with disagreement-as-data reporting; community authority over the evolution of safety definitions made in the community’s name — authority meaning the power to change outcomes, not to advise on them. The stakeholder method adopts value-sensitive design’s machinery (Friedman & Hendry, 2019), indirect stakeholders included: the person affected by the user’s depleted evening, the support worker downstream of the confusing form. Governance architecture — a public methodology commons, community-constitutive standard evolution, revenue arrangements that do not extract community knowledge without return — is a build-before-scale requirement, recorded here as a commitment the program can be held to. The commons has a model in the corpus (§2.11): the bards’ songs reached their audience on home-copied tape, with no owner at the gate; what a community knows about its own safety must travel the same way, without the standard’s owner as gatekeeper.

6.8 Goodhart Resistance

Every audit measure will be optimized against once stakes attach. The program therefore keeps: partially redundant measures per construct (no single score to game); held-out checks not published in advance and rotated per audit cycle; behavioral verification behind self-report (manufactured calm is caught by the veridical/experienced split, §3.6); and periodic

adversarial review in which invited critics try to build interfaces that score well while violating the framework's intent — with successful attacks published and the instrument revised. An audit that cannot survive its own red team should govern no one's interface.

7. The Physiological Hypothesis (Tier C — explicitly quarantined)

This section states the framework's research frontier, separated so that its speculative status cannot contaminate the operational framework, which stands on behavioral evidence alone (§0, §1.3).

The hypothesis. Sustained interaction with dimensionally incoherent, rhythm-blind, or extractive interfaces levies cumulative physiological cost — allostatic load in the predictive-regulation sense (Sterling, 2012; McEwen & Akil, 2020). The body budgets for anticipated demand; uncertainty and vigilance are metabolically expensive; and environments that chronically force expectation revision, sustain unresolved conflict, or compress recovery are hypothesized to tax that budget measurably. This is a claim about demand accounting, not about emotion. It predicts expenditure signatures (autonomic and endocrine markers under consented study), not affective states. It is consistent with, and motivated by, the constructionist position that affect is constructed while bodily budgeting is continuous (Barrett, 2017).

What the hypothesis does not claim. No universal stress response to any stimulus; no polyvagal mechanism; no emotion readable from physiology (an arousal signature is ambiguous among excitement, effort, and distress — Boehner et al., 2007); no dependence on the free energy principle, whose formal status is contested (Bruineberg et al., 2022) and which bears no load anywhere in this framework; no claim that a single session produces clinically meaningful load.

Falsification conditions. The hypothesis is abandoned or restructured if consented ambulatory studies (§6.6) show, with adequate power: no reliable expenditure difference between coherent and incoherent configurations of matched intensity; or expenditure differences that fail to track behavioral strain measures (which would mean the behavioral and physiological constructs dissociate and the framework's unification is wrong); or effect sizes too small to matter at realistic exposure — in which case the framework's harms case rests, honestly, on behavior and experience alone.

Why carry the hypothesis at all. Because if it is true, the harm category changes. Design-induced demand becomes a public-health exposure with an aggregation argument: billions of mildly costly sessions a day are a population-scale exposure in the same formal class as environmental stressors, and regulation of design mechanics acquires the evidentiary structure that environmental-health regulation has. The form of the argument is Tolstoy's

(1869): what no single act can do, the sum of negligible acts does (§2.11). The hypothesis is worth testing because its confirmation would matter and its disconfirmation would discipline this framework's language. Until then it funds no doctrine and licenses no audit finding.

8. Limitations

Validation status. UTDA is not an empirically validated model; the taxonomy is authored, pending derivation (§6.3); every numeric constraint is heuristic; stage-two audit judgments are advisory pending reliability (§6.1). The paper's authority claims are conditional accordingly, and the claim-tier system (§1.3) exists to keep them so.

The taxonomy may be wrong in structure, not only in calibration. The dimensions may fail factor analysis; coherence may not survive the factorial test. The program's design makes such failures informative, but they remain possible, and the framework's continuation is conditional on the result of §6.2.

Cultural generalizability is unresolved. The relevant cognitive and HCI evidence over-samples WEIRD populations (Henrich et al., 2010); until §6.4 reports, all profile ranges are provisional and no configuration should be marketed as universally compatible. The author's own corpus (§2.11) is two canons, not the world, and the same limit applies to it.

The unit-of-analysis limitation. The framework evaluates interface states and sessions. Significant documented harms of current systems are relational and longitudinal — parasocial attachment, dependency, displaced human contact (MIT Media Lab & OpenAI, 2025) — and the dimensional properties of a state cannot carry the full explanatory weight for relationship-level harm. A temporal extension (session, week, relationship horizons) is future work; claims at those horizons are not made here.

The interface boundary problem. A dimensionally impeccable interface to an unjust system remains unjust. A gig-work scheduling application can be coherent, legible, gently paced — and still an instrument of extraction, because the injustice lives in what the system does, not in how it feels. A calm surface over adverse terms is anesthesia, and this framework, applied without its own Doctrine X, could be used to produce it. Gogol's clerk (1842) is handled correctly at every window and destroyed by the building (§2.11); no audit of the windows would have found it. Dimensional evaluation is therefore necessary and not sufficient: it evaluates the environment's demands, not the legitimacy of the arrangement the environment serves, and a conformance claim must never be marketed as ethical clearance of the underlying system. The manufactured-calm provisions (§3.6) police the interface-level version of this failure; the system-level version is beyond this framework's scope and is stated as such.

Adoption physics. The Doctrine conflicts with engagement-optimized business models; a framework without launch-blocking authority becomes theater; short-window experiments disfavor regulation-preserving variants by construction (§5.7). The governance provisions address these; they do not dissolve them.

Authorship and standpoint. This framework was authored by one researcher with professional and lived-experience grounding and, at this stage, without the community accountability structures its own doctrine requires. The participatory track (§6.7) is the remedy; until it has run, the framework's neurodivergence claims should be read as proposals awaiting the only validation that can legitimate them. The corpus declaration (§2.11) makes the author's priors legible; it does not make them representative.

Measurement inheritance. Semantic load inherits the cognitive-load measurement disputes; affective affordance inherits the difficulty of appraisal measurement; the framework's reliability is bounded by its weakest instrument, and §6.1 exists to find out which one that is.

9. Conclusion

Digital systems set the conditions under which people think, choose, feel, recover, and participate, and the institutions around those systems have begun, unevenly and unmistakably, to treat design mechanics as a public concern. This paper has proposed a vocabulary and a research program for that concern: five stimulus-side dimensions anchored to prior constructs; coherence as a falsifiable structural property; predictability split into veridical and experienced forms so that felt calm must be earned by honest behavior; affect handled as affordance rather than delivered state; a doctrine restated in checkable form and extended by a categorical prohibition on turning evaluation into surveillance; pluralism made structural through profiles rather than gestured at in caveats; co-production made constitutive rather than acknowledged; a validation program ordered so that the framework earns authority in the sequence in which it makes claims; and, in this revision, a declared provenance, so that the reader can see where the author's instruments came from.

The central claim survives every revision made here, because it was never a claim about physiology: a digital environment should not manufacture unnecessary strain, hide it, or profit from it. What has changed is the discipline around that claim — what it asserts, at which tier, on what evidence, governed by whom, and formed by what. The standard the paper sets for itself is the one Akhmatova accepted in the prison queue (§2.11): asked whether this could be described, to answer that it can, and then to describe it exactly enough to be checked. If the coherence prediction fails its test, this framework reduces to a well-organized reading of its ancestors, and that result will be published. If it holds, then interface quality has a structural

property that single-axis frameworks cannot see, audit machinery has a construct worth its name, and the environments where people now spend most of their waking attention acquire a design science adequate to what they have become.

Appendix A: Claim Classification Audit (Revised)

Claims are tiered per §1.3; representative examples follow.

Tier A — Behavioral-normative (defensible now). Cancellation flows with asymmetric effort and guilt-laden language obstruct user intent (enforcement-documented). Attention-capture patterns (infinite scroll, autoplay, urgency without referent) measurably extend use beyond intention (Monge Roffarello et al., 2023). Defaults presupposing high sensory tolerance impose differential burden (accessibility and neurodivergence literatures). Interfaces should not be engineered to feel safer than they behave.

Tier B — Theoretical synthesis (judged by fruitfulness, pending derivation). Interface experience can be analyzed through dimensions that distribute cognitive and affective demand. Predictability emerges from dimensional coherence rather than consistency alone. Affective affordance is produced by configuration, not by copy or style alone. Coherence predicts strain beyond intensity (the flagship claim — Tier B until §6.2, then whatever the data says).

Tier C — Physiological hypothesis (quarantined, falsifiable). Sustained dimensional incoherence and extractive demand levy allostatic cost measurable in consented ambulatory study. Carries the falsification conditions of §7; funds no doctrine.

Tier D — Operational proposals (conditional on validation and governance). The audit protocol, state classification, dimensional budgets, generative constraint schema, and certification machinery — advisory until §6.1 reliability, participatory legitimacy per §6.7, and anti-gaming provisions per §6.8.

Outside the tiers — The formative corpus (§2.11, Appendix D). Source of constructs and language; evidential weight nil by declaration.

Appendix B: Glossary (Revised)

Affective affordance — The appraisals an interface makes cheap or costly through its configuration; replaces “emotional valence.” Not a claim about any user’s actual emotional state.

Allostatic load (as used here) — Hypothesized cumulative cost of regulation under environmental demand (Tier C); predictive-regulation sense (Sterling, 2012).

Auro Prism Doctrine — Eleven operational principles for design evaluation and governance, each with a checkable form (§4).

Capability-building demand — Chosen, legible, paced demand whose yield accrues to the user (§3.10); contrast extractive demand.

Dimensional aesthetics — The total perceptual, temporal, semantic, and interactional structure through which an interface is experienced.

Dimensional coherence — The degree to which interface dimensions reinforce rather than contradict one another; operationalized as absence of specified cross-dimensional conflicts (§3.8).

Dimensional profile — Population- and context-calibrated tolerance ranges across the five dimensions; the framework's unit of prescription (§3.9).

Formative corpus — The declared set of literary and musical works from which the author's constructs were formed (§2.11, Appendix D); a source of constructs and language, never of evidence.

Interaction depth — Number, complexity, reversibility, and consequence of actions required to complete or understand a task.

Manufactured calm — Divergence of experienced from veridical predictability: sedative presentational character over opaque or adverse behavior (§3.6). A failure mode, not a compliment.

Motion energy — Quantity, intensity, speed, frequency, and behavioral role of interface movement.

Nervous-system compatibility — Vision term (see Terminology Note): environments whose demands remain within and respectful of human regulatory capacity; technical content = regulatory burden (behavioral) + allostatic hypothesis (Tier C).

Neurodivergent safety (baseline) — A co-produced floor condition in which sensory, cognitive, attentional, and regulatory variance is central rather than exceptional; plural by construction.

Regulatory burden — Total demand an interface places on attention, interpretation, sensory processing, decision-making, and recovery; behavioral construct (Tier A/B), with physiological accounting hypothesized at Tier C.

Semantic load — Interpretive effort required by language, structure, tone, and implication; structural, not word-count.

Sensing Wall — Doctrine XI: dimensional evaluation measures artifacts and consenting study populations, never individual users' physiological or affective states in production.

Temporal rhythm — How an interface structures pacing, interruption, waiting, response, transition, and recovery.

Veridical / experienced predictability — Correspondence of the user's model to actual system behavior (behaviorally assessed) vs. felt stability (self-reported); Doctrine II requires the former to ground the latter.

Visual density — Amount, complexity, and competitive intensity of visible information in an interface state.

Appendix C: Measurement Protocols (Revised)

Ordered by evidentiary role; physiological measures appear last and only under the Sensing Wall.

Behavioral measures (primary): task completion; error rate; reversal/undo frequency; abandonment; hesitation latency before high-consequence actions; repeated scanning; help-seeking; expectation-violation rates (predictability); effort symmetry (entry vs. exit paths); recovery time after interruption; return-with-regret patterns over declared windows.

Self-report measures (paired with behavior, never alone for doctrine findings): perceived cognitive load (with instrument caveats per §2.1); perceived control and self-efficacy (validated instruments, §2.5); perceived clarity; sensory comfort; felt pressure; trust; decision confidence; post-task depletion or relief. The veridical/experienced split (§3.6) requires reporting behavioral and self-report results side by side; divergence is a finding.

Expert review measures: two-stage protocol per §5.3 with inventory/judgment separation; doctrine conformance per checkable forms; conflict inventory; dark-pattern screening mapped to the ontology (Gray et al., 2024); accessibility and neurodivergent-safety review with mixed panels and disagreement reporting.

Physiological measures (Tier C research contexts only, consented, Sensing Wall applies): ambulatory cardiac and electrodermal measures with experience sampling; interpreted exclusively as demand expenditure, never as decoded emotion (Boehner et al., 2007; Barrett et al., 2019); excluded from production systems and product optimization categorically.

Appendix D: The Formative Corpus — Register

Each entry names a work, the construct or section it shaped, and what it supplied. The register is the author's declaration under §2.11; it is not a bibliography of evidence, and every entry carries the same epistemic weight: ancestry, not support. Works are listed in the order of the groups in §2.11.

Akhmatova, Requiem (written 1935–1961; published 1963) — §5.3, §9. The standard of description: asked whether this can be described, to say that it can, and then to describe it exactly enough to be checked.

Solzhenitsyn, "Live Not by Lies" (1974) — §5.7 anti-washing. The minimal ethic of a person inside a system of untruths: do not sign what you know to be false. A conformance claim over an unpublished method is a signed falsehood.

Bykaŭ, the war novels (Sotnikov, 1970, among them) — §5.3 review protocol. The camera held on one person under impossible conditions; the refusal to look away from what the environment does to them.

Orwell, Nineteen Eighty-Four (1949) and "Politics and the English Language" (1946) — §3.5 semantic load; Doctrine XI. Vocabulary as an instrument that prices thought; the telescreen as the device that watches back.

Pavlov, Conditioned Reflexes (1927) — §2.2, §3.8. Science, not literature, and already cited as such; listed here because §2.11 pairs it with its literary double.

Beckett, Waiting for Godot (1952/1954) — §3.3 temporal rhythm. Forced waiting for an arrival the waiting does not control; the two who go on anyway.

Blok, "Night, street, lamp, pharmacy" (1912) — §3.3 temporal rhythm; Doctrine III. The loop with no end state, returning to its own first line.

Dostoevsky, The Idiot (1869) — §2.4 double empathy; §3.6 veridical sincerity. A man who takes words as meant, read as defective by a room that runs on their not being; the failure located in the room.

Gogol, The Overcoat (1842) — §8 interface boundary problem. A clerk handled correctly at every window and destroyed by the building.

Chekhov, the stories (1880–1903) — §5.3; the prose standard of this revision. Say the observable thing and stop.

Dickens, Little Dorrit (1857) — Doctrine IV, effort symmetry. The Circumlocution Office: the classic English description of an obstruction pattern.

Mayakovsky, the agitational verse and ROSTA windows (1919–1922) — §3.7 affective affordance. The poster voice: urgency as a cheap appraisal, the loudspeaker as an interface.

Nabokov, Lectures on Literature (1980) — §5.3 inventory stage; the prose of v5. Precision as an ethic: notice and record the detail before judging it.

Shelley, *Frankenstein* (1818); Bulgakov, *Heart of a Dog* (1925/1987) — Doctrine IX. The maker answers for what it makes, including what it did not review.

Tolstoy, *War and Peace* (1869), the argument on history — §7 aggregation. What no single act does, the sum of negligible acts does.

Martel, *Life of Pi* (2001) — §3.6 manufactured calm. Two accounts of the same events and the question which is the better story; the demand that the better story also be the true one.

Chbosky, *The Perks of Being a Wallflower* (1999); Asher, *Thirteen Reasons Why* (2007) — §3.10, §8; the first-person witness form. The observer at the edge of the room; the chain of small acts that adds up to one outcome.

The bards — Vysotsky, Okudzhava, Galich, Vizbor, Gorodnitsky, Kim, Matveeva, S. and T. Nikitin, Kukin, Dolsky, Mityaev, Rozenbaum, Shaov (recordings, 1950s–2000s) — §6.7 methodology commons; Doctrine XI; the Canon's cadence. Transmission without a gatekeeper; voice before production; Vysotsky's empathy by inhabiting rather than inferring.

The rock of the 1980s–1990s — Tsoi and Kino, Bashlachev, Letov and Grazhdanskaya Oborona, Yanka Dyagileva, Zemfira — the Canon's register. The plain line said in one's own voice; the era's vocabulary of refusal.

Maŭr, *Palessie Robinsons* (1930); Voinovich, *Moscow 2042* (1986) — §3.0, §3.10. The childhood habit of asking what a designed environment does to the people inside it, and what a city built to a plan does to its citizens.

Dick, *Time Out of Joint* (1959); *The Man in the High Castle* (1962); *Ubik* (1969) — §3.6 veridical predictability; Doctrine IX. The environment assembled around a person without their review; authenticity that lives in an object's history, not its surface; the surface that stays calm while the world behind it decays.

Dick, "The Minority Report" (1956); *Do Androids Dream of Electric Sheep?* (1968); *A Scanner Darkly* (1977) — Doctrine XI; §2.4; §3.1. Inferred intent acted on before the act; humanity read off the capillaries; the scanner that unmakes its watcher; kipple.

Dick, "How to Build a Universe That Doesn't Fall Apart Two Days Later" (1978/1985) — §3.6. The definition of reality the audit uses.

Brodsky, Nobel Lecture (1987) — §3.0. The reason the framework keeps the word “aesthetics.”

References

- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of CHI 2019*. ACM. <https://doi.org/10.1145/3290605.3300233>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bardzell, J., Bardzell, S., & Hansen, L. K. (2015). Immodest proposals: Research through design and knowledge. *Proceedings of CHI 2015* (pp. 2093–2102). ACM. <https://doi.org/10.1145/2702123.2702400>
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of FAccT 2021* (pp. 610–623). ACM. <https://doi.org/10.1145/3442188.3445922>
- Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How emotion is made and measured. *International Journal of Human-Computer Studies*, 65(4), 275–291. <https://doi.org/10.1016/j.ijhcs.2006.11.016>
- Bruineberg, J., Dołęga, K., Dewhurst, J., & Baltieri, M. (2022). The emperor’s new Markov blankets. *Behavioral and Brain Sciences*, 45, e183. <https://doi.org/10.1017/S0140525X21002351>
- Calm Tech Institute. (2024). *Calm Tech certification standard*. <https://www.calmtech.institute>
- Calvo, R. A., & Peters, D. (2014). *Positive computing: Technology for wellbeing and human potential*. MIT Press.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Lawrence Erlbaum.
- Case, A. (2015). *Calm technology: Principles and patterns for non-intrusive design*. O’Reilly Media.
- Chen, J. X., McDonald, A., Zou, Y., Tseng, E., Roundy, K. A., Tamersoy, A., Schaub, F., Ristenpart, T., & Dell, N. (2022). Trauma-informed computing: Towards safer technology experiences for all. *Proceedings of CHI 2022* (pp. 1–20). ACM. <https://doi.org/10.1145/3491102.3517475>
- Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.
- Coburn, A., Vartanian, O., & Chatterjee, A. (2017). Buildings, beauty, and the brain: A neuroscience of architectural experience. *Journal of Cognitive Neuroscience*, 29(9), 1521–1531. https://doi.org/10.1162/jocn_a_01146
- Coburn, A., Vartanian, O., Kenett, Y. N., Nadal, M., Hartung, F., Hayn-Leichsenring, G., Navarrete, G., González-Mora, J. L., & Chatterjee, A. (2020). Psychological and neural responses to architectural interiors. *Cortex*, 126, 217–241. <https://doi.org/10.1016/j.cortex.2020.01.009>

- Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press. <https://doi.org/10.7551/mitpress/12255.001.0001>
- Cox, A. L., Gould, S. J. J., Cecchinato, M. E., Iacovides, I., & Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. *CHI EA 2016* (pp. 1389–1397). ACM. <https://doi.org/10.1145/2851581.2892410>
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.
- Ellard, C. (2015). *Places of the heart: The psychogeography of everyday life*. Bellevue Literary Press.
- European Commission. (2026). Preliminary findings on TikTok’s compliance with the Digital Services Act: Addictive design and protection of minors. [Enforcement proceeding; pending final decision.]
- Federal Trade Commission. (2025). *FTC v. Amazon.com, Inc.: Stipulated order on Prime enrollment and cancellation practices*. [Settlement, September 2025.]
- Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann.
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7), 999–1010. <https://doi.org/10.3758/BF03194829>
- Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127–138. <https://doi.org/10.1038/nrn2787>
- Gershman, S. J. (2019). What does the free energy principle tell us about the brain? *arXiv*. <https://arxiv.org/abs/1901.07945>
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. *Proceedings of CHI 2018*. ACM. <https://doi.org/10.1145/3173574.3174108>
- Gray, C. M., Santos, C., Bielova, N., & Mildner, T. (2024). An ontology of dark patterns knowledge: Foundations, definitions, and a pathway for shared knowledge-building. *Proceedings of CHI 2024*. ACM. <https://doi.org/10.1145/3613904.3642436>
- Hallnäs, L., & Redström, J. (2001). Slow technology: Designing for reflection. *Personal and Ubiquitous Computing*, 5, 201–212. <https://doi.org/10.1007/PL00000019>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- International Organization for Standardization. (2019). *ISO 9241-210:2019 — Human-centred design for interactive systems*. ISO.
- Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60(3), 269–298. <https://doi.org/10.1016/j.ijhcs.2003.09.002>
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. <https://doi.org/10.1016/j.cognition.2007.05.006>
- Lyngs, U., Lukoff, K., Slovak, P., Binns, R., Slack, A., Inzlicht, M., Van Kleek, M., & Shadbolt, N. (2019). Self-control in cyberspace: Applying dual systems theory to a review of digital self-control tools. *Proceedings of CHI 2019*. ACM. <https://doi.org/10.1145/3290605.3300361>

- Mack, K., McDonnell, E., Jain, D., Wang, L. L., Froehlich, J. E., & Findlater, L. (2021). What do we mean by "accessibility research"? Proceedings of CHI 2021. ACM. <https://doi.org/10.1145/3411764.3445412>
- Mark, G. (2023). Attention span: A groundbreaking way to restore balance, happiness and productivity. Hanover Square Press.
- Mark, G., Gudith, D., & Klocke, U. (2008). The cost of interrupted work: More speed and stress. Proceedings of CHI 2008 (pp. 107–110). ACM. <https://doi.org/10.1145/1357054.1357072>
- Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), Article 81. <https://doi.org/10.1145/3359183>
- Mathur, A., Mayer, J., & Kshirsagar, M. (2021). What makes a dark pattern... dark? Proceedings of CHI 2021. ACM. <https://doi.org/10.1145/3411764.3445610>
- McEwen, B. S., & Akil, H. (2020). Revisiting the stress concept: Implications for affective disorders. Journal of Neuroscience, 40(1), 12–21. <https://doi.org/10.1523/JNEUROSCI.0733-19.2019>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. <https://doi.org/10.1037/h0043158>
- Milton, D. E. M. (2012). On the ontological status of autism: The double empathy problem. Disability & Society, 27(6), 883–887. <https://doi.org/10.1080/09687599.2012.710008>
- Milton, D., Gurbuz, E., & López, B. (2022). The "double empathy problem": Ten years on. Autism, 26(8), 1901–1903. <https://doi.org/10.1177/13623613221129123>
- Miniukovich, A., & De Angeli, A. (2015). Computation of interface aesthetics. Proceedings of CHI 2015 (pp. 1163–1172). ACM. <https://doi.org/10.1145/2702123.2702575>
- MIT Media Lab & OpenAI. (2025). Investigating affective use and emotional wellbeing on ChatGPT [randomized controlled study and platform analysis]. MIT Media Lab / OpenAI. [Working paper, March 2025.]
- Monge Roffarello, A., Lukoff, K., & De Russis, L. (2023). Defining and identifying attention capture deceptive designs in digital interfaces. Proceedings of CHI 2023. ACM. <https://doi.org/10.1145/3544548.3580729>
- Monge Roffarello, A., De Russis, L., & Lukoff, K. (2025). Digital attention heuristics: Supporting user attention in interface design. ACM Transactions on Computer-Human Interaction. <https://doi.org/10.1145/3725215>
- Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. International Journal of Human-Computer Studies, 68(10), 689–709. <https://doi.org/10.1016/j.ijhcs.2010.05.006>
- Nagamachi, M. (1995). Kansei engineering: A new ergonomic consumer-oriented technology for product development. International Journal of Industrial Ergonomics, 15(1), 3–11. [https://doi.org/10.1016/0169-8141\(94\)00052-5](https://doi.org/10.1016/0169-8141(94)00052-5)
- Norman, D. A. (2013). The design of everyday things (rev. ed.). Basic Books.
- Odom, W., Banks, R., Durrant, A., Kirk, D., & Pierce, J. (2012). Slow technology: Critical reflection and future directions. Proceedings of DIS 2012 (pp. 816–817). ACM. <https://doi.org/10.1145/2317956.2318088>
- Pavlov, I. P. (1927). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex (G. V. Anrep, Trans.). Oxford University Press. [Experimental neurosis: Lecture XVIII.]

- Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology*, 9, 797. <https://doi.org/10.3389/fpsyg.2018.00797>
- Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. Z. (2013). Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. *Proceedings of CHI 2013* (pp. 2049–2058). ACM. <https://doi.org/10.1145/2470654.2481281>
- SAMHSA. (2014). SAMHSA's concept of trauma and guidance for a trauma-informed approach (HHS Publication No. SMA14-4884). U.S. Department of Health and Human Services.
- Scott, C. F., Marcu, G., Anderson, R. E., Newman, M. W., & Schoenebeck, S. (2023). Trauma-informed social media: Towards solutions for reducing and healing online harm. *Proceedings of CHI 2023*. ACM. <https://doi.org/10.1145/3544548.3581512>
- Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press. <https://doi.org/10.1093/oso/9780192845290.001.0001>
- Spiel, K., Frauenberger, C., Keyes, O., & Fitzpatrick, G. (2019). Agency of autistic children in technology research: A critical literature review. *ACM Transactions on Computer-Human Interaction*, 26(6), Article 38. <https://doi.org/10.1145/3344919>
- Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior*, 106(1), 5–15. <https://doi.org/10.1016/j.physbeh.2011.06.004>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Vanden Abeele, M. M. P. (2021). Digital wellbeing as a dynamic construct. *Communication Theory*, 31(4), 932–955. <https://doi.org/10.1093/ct/qtaa024>
- W3C. (2023). *Web Content Accessibility Guidelines (WCAG) 2.2*. World Wide Web Consortium.
- W3C Cognitive and Learning Disabilities Accessibility Task Force. (2021). Making content usable for people with cognitive and learning disabilities. W3C Working Group Note. <https://www.w3.org/TR/coga-usable/>
- Weidinger, L., Uesato, J., Rauh, M., et al. (2022). Taxonomy of risks posed by language models. *Proceedings of FAccT 2022* (pp. 214–229). ACM. <https://doi.org/10.1145/3531146.3533088>
- Weiser, M., & Brown, J. S. (1996). The coming age of calm technology. Xerox PARC.
- Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. *Proceedings of CHI 2007* (pp. 493–502). ACM. <https://doi.org/10.1145/1240624.1240704>
- Literary and musical sources (Appendix D). Cited for ancestry only; no entry below carries evidential weight in this paper.
- Akhmatova, A. (1963). *Реквием [Requiem]*. Written 1935–1961; first published Munich, 1963.
- Asher, J. (2007). Thirteen reasons why. *Razorbill*.
- Beckett, S. (1952/1954). *Waiting for Godot*. Grove Press, 1954. (En attendant Godot, Éditions de Minuit, 1952.)
- Blok, A. (1912). «Ночь, улица, фонарь, аптека». In *Пляски смерти [Dances of death]*.
- Brodsky, J. (1987). Nobel lecture, 8 December 1987. The Nobel Foundation.

- Bulgakov, M. (1925/1987). Собачье сердце [Heart of a dog]. Written 1925; first Soviet publication in Знамя, 1987.
- Букай, В. (1970). Сотнікаў [Sotnikov]. Minsk.
- Chbosky, S. (1999). The perks of being a wallflower. MTV Books / Pocket Books.
- Chekhov, A. P. (1880–1903). Рассказы [Short stories]. Various editions.
- Dick, P. K. (1956). The minority report. Fantastic Universe, January 1956.
- Dick, P. K. (1959). Time out of joint. J. B. Lippincott.
- Dick, P. K. (1962). The man in the high castle. G. P. Putnam's Sons.
- Dick, P. K. (1968). Do androids dream of electric sheep? Doubleday.
- Dick, P. K. (1969). Ubik. Doubleday.
- Dick, P. K. (1977). A scanner darkly. Doubleday.
- Dick, P. K. (1985). How to build a universe that doesn't fall apart two days later (written 1978). In I hope I shall arrive soon. Doubleday.
- Dickens, C. (1857). Little Dorrit. Bradbury & Evans.
- Dostoevsky, F. M. (1869). Идиот [The idiot]. Serialized in Русский вестник, 1868–1869.
- Gogol, N. V. (1842). Шинель [The overcoat]. In Сочинения, vol. 3.
- Martel, Y. (2001). Life of Pi. Knopf Canada.
- Маўр, Ж. (1930). Палескія рабінзоны [Palessie Robinsons]. Minsk.
- Mayakovsky, V. V. (1919–1922). Окна РОСТА [ROSTA windows] and agitational verse. Various editions.
- Nabokov, V. (1980). Lectures on literature (F. Bowers, Ed.). Harcourt Brace Jovanovich.
- Orwell, G. (1946). Politics and the English language. Horizon, 13(76), 252–265.
- Orwell, G. (1949). Nineteen eighty-four. Secker & Warburg.
- Shelley, M. (1818). Frankenstein; or, The modern Prometheus. Lackington, Hughes, Harding, Mavor & Jones.
- Solzhenitsyn, A. (1974). Жить не по лжи! [Live not by lies]. Samizdat essay, 12 February 1974.
- Tolstoy, L. N. (1869). Война и мир [War and peace]. Serialized in Русский вестник, 1865–1869; the argument on history in the Epilogue, part 2.
- Voinovich, V. (1986). Москва 2042 [Moscow 2042]. Ardis.
- Авторская песня (bard song), recordings and magnetized copies, 1950s–2000s: В. Высоцкий, Б. Окуджава, А. Галич, Ю. Визбор, А. Городницкий, Ю. Ким, Н. Матвеева, С. и Т. Никитины, Ю. Кукин, А. Дольский, О. Митяев, А. Розенбаум, Т. Шаов. [Corpus entry; discography not itemized.]
- Русский рок 1980-х–1990-х: Кино (В. Цой), А. Башлачёв, Гражданская оборона (Е. Летов), Янка Дягилева, Земфира. [Corpus entry; discography not itemized.]
- Citation integration note (carried from v3, updated for v5): DOIs and publisher details require verification against a citation manager before external submission. The European Commission (2026), FTC (2025), and MIT Media Lab & OpenAI (2025) entries reference enforcement records and a working paper whose canonical citation forms should be confirmed at submission time. The Monge Roffarello et al. (2025) TOCHI title should be verified against the published version. The literary and musical entries

are corpus declarations: editions and first-publication details should be completed from the author's own copies before submission, and the two collective entries may be itemized or left as corpus entries at the author's discretion.

ADDENDUM — The Canon Edition Mapping

The academic paper above is the core. The Auro Prism Canon edition — the charter document in the voice of Volumes I–III — carries everything above plus the layers a journal cannot hold. This addendum specifies what the Canon edition adds, so both texts can be maintained from this single source without drift.

Voice and register. The Canon edition restores the doctrine's full rhetorical register — the short-line cadence of Volumes I–III — for the doctrine statements themselves, while keeping this paper's checkable forms as each doctrine's "operational face." Format: Canon voice states the principle; a boxed operational clause carries the checkable form verbatim from §4. The two must never diverge in substance; the checkable forms are canonical.

Volume linkages. Each doctrine cross-references its Canon ancestry: Doctrines I–VIII to Volume II's First Principles and Sacred Constraints (dignity before optimization; agency before persuasion; transparency before manipulation); Doctrine IX to Volume II's AI Doctrine chapters (explainability, accountability, reversibility, human authority); Doctrine X to "Ethics Is Structural" and Volume II's Red Lines; Doctrine XI is new to the Canon and enters Volume II's Sacred Constraints as a red line ("What Auro Prism Refuses" gains: inferring the inner states of the people it serves). The Pulse framework (Volume III, Part I) is annotated as the organizational-scale sibling of the dimensional audit — Pulse listens at the level of institutions to what the audit measures at the level of interfaces — and Pulse's Sacred/Structural/Operational layering is mapped to the claim tiers (Sacred $\approx$ Tier A normative commitments; Structural $\approx$ Tier D governance; Operational $\approx$ audit practice).

Pattern library integration. Volume III's thirteen anti-patterns are annotated with their dimensional signatures — e.g., Care Theater as an affective-affordance/veridicality divergence (the organizational form of manufactured calm); False Alignment as coherence failure at the social layer; Explanation Theater as the transparency analogue of manufactured calm (Doctrine VIII's named enemy); Captive Loyalty and Forced Intimacy as extractive-demand patterns at relationship scale, flagging the unit-of-analysis limitation (§8) as Canon future work. The Canon edition may make these linkages richly; the academic paper deliberately does not claim them.

The corpus in the Canon. The formative corpus (§2.11, Appendix D) is Canon material. The paper cites it for ancestry in one sentence at a time; the Canon edition and the Field Notes may quote it, argue with it, and write in its cadence. The same boundary holds in both

directions: the Canon may use the corpus for voice, never for claim strength, and no checkable form acquires literary content in either edition.

What the Canon edition keeps that the paper omits. The meta-principle in full ("We design systems worthy of human beings."); the Steward's Oath as the practitioner's commitment ceremony for audit training; the "future is a stakeholder" horizon (multi-lifespan framing) as the Canon's answer to the paper's product-cycle scope; worked reflections and repair pathways in the Volume III style.

What the Canon edition must also carry from this revision — no exceptions. The Terminology Note (§0) in full; the claim-tier system; the Sensing Wall; the co-production governance commitments; the anti-washing provisions; and the corpus boundary stated above. The Canon's moral register heightens rather than loosens the epistemic discipline: doctrine without the tier system is exactly the "ideology to buyers, vibes to scientists" failure the review warned against, and the Canon edition's credibility depends on carrying its own constraints visibly.

Maintenance rule. The academic core is the source of truth for constructs, claims, and checkable forms. Canon prose may exceed it in voice, never in claim strength. When validation results (§6) revise a construct, both editions version together, and the Summary of Revisions genre established in v4 is the changelog mechanism for both.

End of Revised Draft v5.